ice Level easurement:

Proceedings from a conference sponsored by Statistics Canada

La mesure du niveau des prix:

Actes du colloque tenu sous l'égide de Statistique Canada

Canada

Price Level Measurement: La mesure du niveau des prix:

Proceedings from a conference sponsored by Statistics Canada

Actes du colloque tenu sous l'égide de Statistique Canada

Editors:

W.E. Diewert
C. Montmarquette

Éditeurs:

W.E. Diewert
C. Montmarquette

Published under the authority of the Minister of Supply and Services Canada

Publication autorisée par le ministre des Approvisionnements et Services Canada

The views expressed by the authors are their own and not necessarily those of Statistics Canada

Les opinions exprimées dans ce rapport sont celles des auteurs et ne reflètent pas nécessairement celles de Statistique Canada

December 1983
62-x-502

Price: $50.00

ISBN 0-660-52186-5

Ottawa

Décembre 1983
62-x-502

Prix: $50.00

ISBN 0-660-52186-5

Ottawa

FOREWORD

In November 1982 in Ottawa, Statistics Canada hosted a conference on price level measurement. More than 350 people took part in presenting the 50 papers and commentaries on the agenda. As explained by C. Baumgarten and C.D. Hodgins in the introduction to this volume, the aim of the meeting was to analyse the concepts and methodologies involved in price indexes, especially the Consumer Price Index.

The introduction also contains a summary, prepared by W.E. Diewert, of the principal ideas expressed in the papers and commentaries selected for inclusion in this publication. In keeping with the federal government's language policy, a number of the papers are presented in Canada's two official languages.

The conference and this volume would not have been possible without the assistance of Mr. M.B. Wilk, Chief Statistician of Canada. The Research Program Committee, composed of W.E. Diewert (University of British Columbia), C. Montmarquette (University of Montreal), C.D. Hodgins (Western Economic Services Limited), C. Baumgarten and B.J. Szulc (Statistics Canada), is very grateful to Mr. Wilk and the members of The Price Level Measurement Review Program Committee for their support and encouragement.

Cynthia Baumgarten deserves special mention for her helpfulness and administrative expertise, and Erwin Diewert's intellectual leadership on the Committee was appreciated by all concerned.

The Committee would like to thank everyone who contributed to the success of the Price Measurement Review Program.

Claude Montmarquette

AVANT-PROPOS

Statistique Canada tenait à Ottawa, en novembre 1982, une conférence sur la mesure du niveau des prix. Plus de 350 personnes ont participé à la présentation des quelque cinquante textes et commentaires inscrits au programme. Cette rencontre, comme l'expliquent C. Baumgarten et C.D. Hodgins dans l'introduction du présent volume, visait à permettre l'analyse des concepts et méthodologies liés aux indices de prix et, en particulier, l'Indice des prix à la consommation.

On trouvera également en introduction une synthèse, rédigée par W.E. Diewert, des principales idées contenues dans les textes et commentaires que nous avons sélectionnés afin de les regrouper dans ce volume. On notera que plusieurs textes sont publiés dans les deux langues officielles du Canada, en conformité avec la politique linguistique du gouvernement fédéral.

La conférence et le présent volume n'auraient pu voir le jour sans l'appui de Monsieur M. Wilk, statisticien en chef du Canada. Le Comité responsable du programme de recherche composé de W.E. Diewert de l'Université de la Colombie-Britannique, C. Montmarquette de l'Université de Montréal, C.D. Hodgins de Western Economic Services Limited, C. Baumgarten et B.J. Szulc de Statistique Canada, est très reconnaissant envers le Docteur Wilk et les membres du Comité sur la mesure du niveau des prix, pour leur support et leurs encouragements.

La grande disponibilité et la compétence administrative de Cynthia Baumgarten sont à souligner et tous ont apprécié, au sein du Comité responsable, le leadership intellectuel d'Erwin Diewert.

Le Comité responsable remercie de plus, tous les participants qui ont contribué à la réussite du programme de l'examen de la mesure des prix.

Claude Montmarquette

TABLE OF CONTENTS — TABLE DES MATIÈRES

TABLE OF CONTENTS (Continued) — TABLE DES MATIÈRES (Suite)

Page

Section II (Concluded) — (Fin)

The Theoretical Foundations of a Consumer Price Index
Les fondements théoriques d'un Indice des prix à la consommation

Section III

Specific Price Measurement Issues
Problèmes particuliers de la mesure de la variation des prix

TABLE OF CONTENTS (Continued) — TABLE DES MATIÈRES (Suite)

Page

Section III (Concluded) — (Fin)

Specific Price Measurement Issues
Problèmes particuliers de la mesure de la variation des prix

TABLE OF CONTENTS (Concluded) — TABLE DES MATIÈRES (Fin)

SECTION I

Introduction

REMARKS OPENING THE CONFERENCE ON PRICE LEVEL MEASUREMENT

Pour vous fournir une version dans la langue officielle de votre choix, le texte anglais est suivi du texte français (p.14) dans cette publication.

Martin B. Wilk
Chief Statistician of Canada

On behalf of Statistics Canada, I want to welcome you to this Conference on Price Level Measurement. Our objective in this Conference is to assure that the price measurement information Statistics Canada provides is developed as effectively and as appropriately as possible.

There is general recognition that inflation is a crucial economic characteristic. Historically, inflation has had enormous influence on the political and social structure of societies, from time to time.

In recent years inflation has been one of the foremost targets of monetary and fiscal policy in many countries of the world, including both Canada and the United States. Inflation has been a major preoccupation of World Credit Markets, and a central issue in wage and salary negotiations. Theories, policies and practices for reducing, or coping with, inflation have been the subject of much attention and controversy.

In view of the importance of inflation, and of the controversies regarding its causes and control, it is reasonable that statistical organizations which produce measures of inflation, have at times been challenged regarding the soundness of their concepts and methods. Some critics have even seemed to imply that the measurement index itself is a major cause of inflation. Nonetheless, challenges and criticisms play a key role in scientific and statistical work, and Statistics Canada pays close attention to all serious criticisms and suggestions regarding the validity and credibility of the statistics for which it is responsible.

Statistics Canada has an extensive program of price level measurement and produces a number of measures of inflation. We have statistics which show the change in the average level of prices for manufactured goods for energy, for transportation, for prices received by farmers, and so on. We produce implicit price indexes derived from current and constant dollar measures of the components of gross national expenditure. And we also produce the Consumer Price Index.

Of these many measures, the Consumer Price Index has, for various reasons, been embraced as **the** measure of the rate of inflation in the minds of most Canadians. It is commonly referred to as the cost-of-living index, and is widely used in numerous contexts in efforts to protect the real purchasing power of many kinds of payments, such as wages and pensions. The range and types of uses to which the CPI is put surely go well beyond whatever **any** index could possibly satisfy perfectly, whatever its definition.

In recent years, concerns have been expressed by representatives of labour unions, by consumer associations, by the business sector, by government, by various economic analysts and by journalists as to the validity of the Consumer Price Index as the principal measure of inflation in Canada or in some of the regions. In addition, there have evolved some misunderstandings as to exactly what the Consumer Price Index is, what it is measuring and what it is not measuring. Some concerns about the CPI reflect the point of view that it should be a cost-of-living index. Some accept it as a cost-of-living index in principal but claim that it overstates it in practice because it fails to take adequate account of substitutions that consumers make in the market place, as they cut back on those products whose relative price has increased, and expand purchases of other goods and services whose relative price has declined. Others have argued that the CPI understates inflation experienced by the poor or the elderly or other groups.

While the rate of change in the Consumer Price Index has slowed considerably during 1982, in Canada and in the United States, inflation remains a high priority for economic policy makers and the public. So long as this remains the case, concerns about the quality of inflation data may be expected to continue. Criticisms of the data that are not adequately dealt with could undermine the credibility of our indexes and add another element of confusion and controversy to public policy debate.

Over the years, Statistics Canada has maintained an ongoing research and development program to ensure that concepts and methods are suitably adapted to changing circumstances, so as to produce measures of price movement, including the CPI, of the highest feasible quality.

The staging of this Conference is not to suggest that we have discovered, or have had brought to our attention, something fundamentally wrong with the Consumer Price Index. Rather, it is because of the enormous importance rightly attributed to measures of inflation that we undertook, over a year ago, to augment our in-house studies with supplementary initiatives to review price measurement, culminating in this public Conference and the publication of its proceedings.

Our review program comprised several important elements. We solicited views from a number of prominent economists and researchers. We consulted with major users of price data in a series of meetings in every province and territory in Canada. Input was received from provincial and territorial governments, labour and business groups, the media and many others. The results of these meetings were used to give direction to special studies by leading professional economists.

The organization of this research work was the responsibility of a Price Level Measurement Research Program Committee, under the co-chairmanship of Professor Erwin Diewert of the University of British Columbia and Professor Claude Montmarquette of the Centre de Recherche en Développement Économique of the University of Montreal. Other committee members are Dr. Cyril D. Hodgins, President of Western Economic Services Ltd., of Vancouver, B.C., who also played a major role in the consultations, and two members of Statistics Canada's staff, Mr. Bohdan Szulc, Chief of Research of our Prices Division and Mrs. Cynthia Baumgarten, Manager of the Price Measurement Review Program.

I want to extend my sincere thanks to the Program Committee. They have done an outstanding job, in planning and organizing this Conference, involving some of the world's leading price measurement specialists and economists. There was also a substantial contribution to this program by my collegues Guy Leclerc, Stu Wells, Bernie Lynch and Denis Desjardins.

It is with great pleasure that I now introduce the co-chairman of the Program Committee, Dr. Erwin Diewert, and invite him to review the Conference agenda with you.

DISCOURS D'OUVERTURE DE LA CONFÉRENCE SUR LA MESURE DU NIVEAU DES PRIX

To provide you with a version in the official language of your choice, the French text is preceeded by the English text (p.11) in this publication .

Martin B. Wilk,
Statisticien en chef du Canada

Au nom de Statistique Canada, je désire vous souhaiter la bienvenue à cette Conférence sur la mesure du niveau des prix. Nous voulons par cette conférence nous assurer que les renseignements de Statistique Canada sur la mesure du niveau des prix sont élaborés de la façon la plus efficace et la plus appropriée.

Tous reconnaissent que l'inflation est un indicateur économique crucial. Historiquement, ce phénomène a exercé périodiquement une énorme influence sur la structure politique et sociale des sociétés.

Ces dernières années, l'inflation a été la cible première des politiques monétaires et fiscales dans de nombreux pays, y compris au Canada et aux États-Unis. Elle a été une source de préoccupation sur les marchés mondiaux du crédit et une question au centre des négociations salariales. Les théories, les politiques et les pratiques pour combattre l'inflation ont suscité beaucoup d'attention et de controverses.

Puisque l'inflation est si importante et que personne ne s'entend sur ses causes et sur la façon de la contrôler, il n'est pas étonnant que les concepts et les méthodes utilisés par les organismes statistiques qui produisent des mesures de l'inflation ont parfois été remis en question. Certains critiques ont même insinué que l'indice lui-même était l'une des principales causes de l'inflation. Pourtant, les défis et les critiques jouent un rôle important dans le cadre des activités scientifiques et statistiques, et Statistique Canada prête une attention toute particulière aux critiques et aux suggestions formulées sur la validité et la crédibilité de ses statistiques.

En plus de son programme élaboré de mesure du niveau des prix, Statistique Canada produit un certain nombre de mesures de l'inflation. Ses statistiques montrent la variation des prix moyens des produits manufacturés, de l'énergie et du transport, des prix payés aux agriculteurs, et ainsi de suite. Il produit aussi, outre l'indice des prix à la consommation, des indices implicites de prix à partir de composantes, mesurées en dollars courants et en dollars constants, de la dépense national brute.

Pour diverses raisons, l'indice des prix à la consommation est la principale mesure du taux d'inflation dans l'esprit de la plupart des Canadiens. On l'assimile souvent à un indice du coût de la vie et on l'utilise beaucoup dans diverses circonstances pour tenter de protéger le pouvoir d'achat réel de différents paiements, tels que les pensions et les salaires. L'IPC, pas plus que n'importe quel **autre** indice, quelle qu'en soit la définition, ne peut se prêter parfaitement à un tel éventail et à une telle diversité d'applications.

Ces dernières années, des analystes économiques, des journalistes ainsi que des représentants des syndicats, des associations de consommateurs, du monde des affaires et de l'administration publique ont exprimé des réserves quant à la validité de l'indice des prix à la consommation comme principale mesure de l'inflation au Canada ou dans certaines régions. En outre, on en est venu à ne plus très bien comprendre la nature exacte de l'indice des prix à la consommation, c'est-à-dire ce qu'il mesure ou ce qu'il ne mesure pas. Certains voudraient que l'IPC devienne un indice du coût de la vie. D'autres pensent qu'il l'est déjà en principe mais qu'en pratique il surestime le coût de la vie en ne tenant pas compte adéquatement des substitutions du consommateur qui, devant l'augmentation du prix relatif de certains produits et services, les remplace par d'autres dont le prix relatif a baissé. D'autres encore affirment que l'IPC minimise les conséquences de l'inflation pour les pauvres, les personnes âgées ou d'autres groupes.

Même si l'indice des prix à la consommation a considérablement ralenti sa progression en 1982, tant au Canada qu'aux États-Unis, l'inflation s'inscrit toujours parmi les préoccupations premières des législateurs et du public. Tant qu'il en sera ainsi, il faut s'attendre à ce que la qualité des données sur l'inflation soit mise en doute. Nous nous devons de réagir adéquatement aux critiques pour préserver la crédibilité de nos indices et pour ne pas ajouter un élément de confusion ou de controverse au débat politique public.

Au fil des années, Statistique Canada a maintenu un programme permanent de recherche et de développement afin que les concepts et les méthodes restent adaptés aux circonstances changeantes et que les mesures des variations de prix, y compris l'IPC, soient toujours de la plus haute qualité possible.

Si nous avons organisé cette conférence, ce n'est pas parce que nous avons découvert, ou que quelqu'un nous a signalé, une lacune fondamentale de l'indice des prix à la consommation. C'est plutôt parce que les mesures de l'inflation ont pris, à juste titre, une très grande importance que nous avons entrepris, il y a plus d'un an, diverses initiatives pour compléter nos études internes sur la mesure de la variation des prix, dont la présente conférence, et la publication de nos délibérations, constituent les points culminants.

Notre programme de révision comprenait plusieurs éléments importants. Nous avons demandé l'avis d'un certain nombre d'économistes et de chercheurs éminents. Nous avons rencontré les principaux utilisateurs des données sur les prix dans chaque province et territoire du Canada. Les administrations publiques provinciales et territoriales, les syndicats et le milieu des affaires, les médias et bien d'autres nous ont aidé. Les résultats de ces rencontres ont permis d'orienter les études spéciales entreprises par des économistes de premier plan.

L'organisation des travaux de recherche a été confiée au Comité du programme de recherche sur la mesure du niveau des prix, coprésidé par M. Erwin Diewert de l'Université de la Colombie-Britannique et M. Claude Montmarquette du Centre de recherche en Développement Économique de l'Université de Montréal. Ont également siégé au Comité, M. Cyril D. Hodgins, président de Western Economic Services Ltd. de Vancouver (C.-B.), qui a également joué un rôle de premier plan dans les consultations, et deux membres de Statistique Canada, M. Bohdan Szulc, chef de la recherche à la Division des prix, et Mme Cynthia Baumgarten, chef du Programme de révision de la mesure du niveau des prix.

Je désire remercier sincèrement les membres du Comité du programme. Ils ont accompli une tâche remarquable en planifiant et en organisant cette conférence, à laquelle participent des économistes et des spécialistes de la mesure du niveau des prix de renommée mondiale. Il y eut aussi une contribution remarquable à ce programme de la part de plusieurs

de mes collègues, dont Guy Leclerc, Stu Wells, Bernie Lynch et Denis Desjardins.

Il me fait grand plaisir de vous présenter maintenant le coprésident du Comité du programme, M. Erwin Diewert, et de l'inviter à passer en revue avec vous l'ordre du jour de la Conférence.

PRICE MEASUREMENT REVIEW PROGRAM: CONSULTATIONS FEEDBACK REPORT

Pour vous fournir une version dans la langue officielle de votre choix, le texte anglais est suivi du texte français (p.38) dans cette publication.

Cynthia Baumgarten
Price Measurement Review Program, Statistics Canada
Cyril D. Hodgins
Western Economic Service Ltd.

I. Introduction

In recent years a number of concerns have been publicly expressed by representatives of labour unions, consumer associations, the retail sector, government and various economic theorists and journalists as to the validity of the Consumer Price Index as the principal measure of inflation in Canada. In addition there has been some misunderstanding as to exactly what the Consumer Price Index is, what it is measuring and what it is not measuring. In the climate of inflation such as now prevails these concerns may be expected to continue and they could undermine the usefulness of the index as a widely accepted measure of retail price changes.

In recognition of such concerns Statistics Canada decided to undertake a program of review, evaluation and information. That program has as its objectives:

(a) to expose the Consumer Price Index to public scrutiny;

(b) to provide a mechanism for in-depth analyses of alternate concepts and methodologies for Consumer Price Index construction. These studies will either confirm present practices as the most effective available or suggest concrete ways in which they might be improved;

(c) to generate greater public understanding of the concepts of aggregate price measurement, thereby helping to lessen confusion in the minds of critics and of the public in general as to what the CPI is, what it realistically can be expected to measure, and what it cannot be expected to measure.

Three principal elements are contained within the review program. They are:

(1) Information Exchange Project
(2) User Consultation Project
(3) Conference Project

The Information Exchange Project was a series of panel discussions arranged as a means whereby Canadian and foreign experts could come together to examine and discuss various aspects of price level measurement and of the inflationary process.

The User Consultation Project consisted of a series of meetings between representatives of Statistics Canada and a wide variety of users of the CPI. The purpose of the meetings was to review the Statistics Canada's program of work in the area of price measurement and to be informed by the participants as to their comments, criticisms and grievances about existing measures. The program of new research to be supported by Statistics Canada subsequently would be guided by the results of these meetings.

The Conference Project will consist of the convening of a public conference at which the results of commissioned research along with other research findings will be presented. It is intended that the papers, discussants' comments and a summary of the conference proceedings be published.

Meetings for the User Consultation phase of the program were arranged through the User Advisory Services representatives at the Statistics Canada Regional Offices and through the provincial and territorial government statistical focal points. Their knowledge of the principal price index users in the regions and within the respective governments, ensured that the invitation to participate was extended to a broad range of individuals and organizations who potentially would have an interest in the program. The final selection of whom to meet with was, in the end, an arbitrary process. Any other procedure, other than the

holding of formal public hearings, also would have been arbitrary. There was neither time nor funding available for the holding of public hearings. Nonetheless, the approach that was adopted did make a positive contribution to opening up a very important part of the statistical process through dialogue with interested users.

The purpose of this report is to summarize the principal findings that emerged from the user consultation meetings.

II. The Consultation Meetings: Content and Coverage

Meetings were held in all 10 provinces and in the Yukon and Northwest Territories. In total, 28 meetings were held in 14 cities during the period January 25, 1982 - March 18, 1982.

The number of attendees in total, by type of organization or affiliation, is shown in Table 1. Over one-half of those consulted were employed by provincial government departments and agencies. The next largest group was the business community, followed by organized labour. The "other" category includes groups such as the Consumers Association of Canada. Media contact was minimal but not for any lack of effort on the part of Statistics Canada. The number of media representatives that attended was only a small portion of those that had been advised of the meetings and invited to attend.

A short, informal presentation by Statistics Canada officials was given to open each meeting. While the presentation varied slightly from meeting to meeting in its points of emphasis, mainly to highlight local problems, issues or concerns, the main thrust was as follows.

Inflation is the foremost target of current monetary and fiscal policy, in Canada and the United States. It is the foremost preoccupation of North American credit markets. Inflation is also the reason why government budget deficits in Canada and the United States are feared by so many, because of their anticipated downstream inflationary consequences. But what is this thing called inflation?

TABLE 1. Attendance at User Consultation Meetings Statistics Canada Price Measurement Review Program

Type of Organization	No. of Attendees	Percent
Organized Labour	16	7
Business	46	20
Media	3	1
Federal Government*	10	4
Provincial Governments:		
Nfld.	6	3
N.S.	4	2
N.B.	13	6
P.E.I.	13	6
Que.	18	7
Ont.	15	6
Man.	10	4
Sask.	7	3
Alta.	16	7
B.C.	14	6
Yukon	11	5
N.W.T.	13	6
Total, Provincial Governments	140	60
Other	19	8
Total Attendees	234	100

* Does not include Statistics Canada representatives.

Any standard textbook definition of inflation describes it as a sustained rise in the general or overall level of prices. It is sustained in the sense of continuing, as opposed to occasional or periodic changes. It involves the general level of prices as compared to relative price adjustments which reflect changes in market demand and supply conditions bearing on specific goods or services, such as oranges, ladies' footwear or gasoline.

There are a number of measures of the rate of inflation. There are indexes which show the change in the average level of prices charged by industry and other indexes which summarize prices received by farmers. There are implicit price indexes derived from the current and constant dollar measures of the components of gross national expenditure.

There is also, of course, the Consumer Price Index, the series which in the minds of most Canadians is **the** rate of inflation. The CPI gained its status as the definitive measure of the rate of inflation for a number of reasons:

- The CPI describes what is happening in general to the prices of those goods and services which we, as consumers, are acquiring from day to day in the market place.
- The CPI is available on a monthly basis and with only a short lag following the end of each month.
- The CPI can be broken down into numerous component parts, so that movements in the total index can be explained in terms of changes in specific components.
- The CPI is available for each of the major urban centres throughout Canada.

Alternative measures of inflation, such as the implicit GNE price deflators, are not nearly so widely watched or used. They do not appear to be generally understood and are available only quarterly. The implicit deflators from the national income and expenditure accounts cannot be as thoroughly disaggregated and are available only on a national level. And so the CPI has emerged as **the** measure of the rate of inflation in the minds of most Canadians. More than that, it is widely referred to as the measure of changes in the **cost of living**, and as such, has come to be widely used in numerous contexts to protect the real or effective purchasing power of a variety of payments. This has been done through the indexation of wages and salaries under collective agreements. It has been done through the indexation of government transfer payments such as family allowances and old-age security payments. As well, it has been used as the trigger for making automatic adjustments in the personal income tax structure with the intention of denying to government the taxation revenue windfall that comes from pushing inflated incomes up through progressive tax rates.

The importance of the CPI places it foremost among all of the economic performance indicators produced by Statistics Canada. The regional offices of Statistics Canada report more enquiries about CPI than about any other single statistical measure.

The importance of the CPI, and the range and types of uses to which it is being put, undoubtedly go well beyond anything the orginal developers of the measure could ever

have imagined. The CPI has been calculated in Canada from as far back as 1913 and up until 1949, was referred to officially as the Cost-of-Living Index.

Most users of the index are aware of what the CPI is intended to be, namely, "The percentage change through time in the cost of a constant basket of goods and services, representing the purchases made by a particular population group in a specified time period". Because it is intended that the basket contain a set of goods or services of unchanging or comparable quantity and quality for which price changes over time are measureable, changes in its cost should be strictly due to price movement.

While that particular description has a bland ring of professional neutrality about it, the measure as it is now calculated is not without its critics. Many concerns about the CPI are well-known and frequently reflect the point of view that it should be more of a cost-of-living index than it in fact is. Some claim that it **overstates** the rate of change in the cost of living because it fails to take account of substitutions that consumers make in the marketplace, cutting back on those products whose relative price has increased, and increasing purchases of those for which prices have declined in relative terms. Others claim that it **understates** the rate of change in the cost of living because it gives too little relative importance in the basket to recently high-flying components as such energy and housing.

While to some persons the CPI may stand for the **Confusing** Price Index or even the **Confounded** Price Index, it is apparent that the majority of Canadians accept what it is telling us about the changing level of prices, perhaps because we really do not have any better alternative. There is nothing else with which it can be directly compared.

Over the years Statistics Canada has maintained an in-house, ongoing research and development program to ensure that concepts and methods are adapted to changing circumstances. However, because of the enormous current importance of the CPI and the extra attention being paid to it, the Chief Statistician decided to augment the in-house program of research in 1982 by launching a new Price Measurement Review Program. It has three principal components:

- a series of bilateral consultations with major users in all regions of Canada;

- Statistics Canada professional and financial support for a group of special studies to be done by leading economists in the academic community;
- a national conference on Price Level Measurement.

The entire approach reflects a healthy open attitude toward the measurement process, and the recognition that such complex and often confusing issues can stand to benefit from an open public exchange of professional views. It does not reflect any recent new found concern that there is something fundamentally wrong with the CPI.

Feedback from the user consultation meetings indicated that this Statistics Canada initiative was being extremely well-received and widely supported throughout Canada. A substantial number of leading economists have indicated a preparedness to participate through carrying out research studies.

Numerous groups indicated a willingness to become involved, through articulating any concerns they may have about the CPI and the way it is calculated, and expressing an interest to participate in the conference itself. With that kind of interest and support we are confident of the program's ultimate success.

It is hoped that the success of this program will encourage Statistics Canada to follow up with similar programs of sponsored research leading up to public conferences on other major statistical programs. In this way, the initiative of a particular federal agency can set an example of openness and public accountability that might be followed by other government departments and agencies.

The following sections summarize some of the feedback that was received from users.

III. Price Level Measurement Issues

This section summarizes the principal areas of concern identified by users during the user consultation process. The order in which they are presented does not reflect any particular priority or urgency. In fact, the priority attached to various concerns varied considerably in accordance with the particular interests of attendees and the area of Canada in which the meeting was being held.

(a) Frequency of Updates for Expenditure Weights

In April 1982, a weighting pattern reflecting 1978 consumer expenditure patterns replaced the 1974 expenditure pattern in use until that time as the basis for determining the relative importance of consumer purchases. Updating the weights from a 1974 base to a 1978 base was viewed as only a partial improvement. Repeatedly, participants at the meetings expressed concern that the weights are ''not current enough''. Specifically, since 1978, consumers have been faced with escalating costs for energy and for products of energy-intensive production processes. It was felt that, in response to these increases, consumer expenditure patterns have changed. The effect of these changes, we were told, is not reflected in the present CPI even after the April 1982 update. The general consensus was that issues such as this, namely out of date weights, undermine the credibility of the CPI.

Even though a more current expenditure pattern could well be only a ''cosmetic'' improvement it nonetheless was considered to be worth examination. Most participants felt that the credibility of the index would be enhanced if there were some non-arbitrary way to update the weighting pattern in times when the component distributions within the family budget appear to be changing significantly. The possibility of estimating weighting diagrams based on other more current indicators which are relevant and available was often suggested as a solution.

As one of the main building blocks for the CPI, the Family Expenditure Survey was thought to be a potential source for improvement in the timeliness of the CPI weights. An overall examination of the concepts and methodology of the survey in light of the needs of the CPI was encouraged.

(b) Shelter Costs

Issues concerning the shelter component of the CPI came up in each city. There was an expressed perception that the homeownership and rental sub-indexes, and especially the rental component, did not reflect the ''reality'' of the market place. Price increases for homeownership and for rented accommodation were thought to be much higher than what was being indicated by the respective CPI sub-indexes. The Royal Trust housing price index and CMHC rental data contributed to that perception. Both of these sources, we

were told, show larger increases over time than the comparable CPI components at city levels.

The current methodology of applying a fixed distribution of outstanding mortgages by age over a 60-month period was thought to cause an understatement of the rate of change of homeownership user-costs, and a delay in reflecting the impact of mortgage interest rate changes. Recently in the marketplace, much shorter terms have been adopted for mortgage contracts and some banks were even offering monthly ''floating'' rate mortgages. Purchasers of new homes are frequently being offered discounted (lower than market) mortgage interest rates for homes whose selling prices include a margin to cover the cost of discounting the mortgage interest rate. Questions as to how widespread these practices are, and what effect they would have on the rate of change of the home ownership component of the CPI (if the methodology were altered to reflect these practices), were presented as possible directions for research.

Another question concerning the shelter component focused on the present formula for determining the user-cost of owned accommodation. It was suggested that the current practice of not weighing the benefits of capital gains against the nominal interest cost component of mortgage payments overstates the ''real cost'' changes that face homeowners.

(c) The Treatment of Taxes and Public Goods

Because the payment of income taxes does not relate to any **specific** quantity of good or service provided to the population, they are excluded from the CPI. Many participants felt that, while this may be an appropriate approach to take in defining a consumer price index for some purposes, it was not appropriate for the official CPI given that it is used as though it were a true cost-of-living index. They observed that the payment of income taxes can take a substantial portion of a family's income and therefore is a very important factor in the cost of living. According to this view, in light of the CPI's widespread application in the indexing of both earned incomes and transfer payments, a more appropriate index is needed to properly reflect the cost of living.

There were suggestions that Statistics Canada develop a system of indexes some of which, when taken together, would approximate a cost-of-living index. The CPI in its present

form would be a component within the system.

"Public goods" (goods and services which are provided by governments for public consumption) are financed by means of the tax system. Although all of the revenues from the tax system may go toward the production of public goods, not all public goods produced are consumed. It therefore could be argued that the CPI should reflect the cost of public goods **consumed** rather than the cost of public goods **produced**. Thus the inclusion of taxes in an index designed to reflect the current cost of living might well overstate (or give too large a weight to) the public goods component.

Except for public goods financed from property taxes, the CPI excludes public goods from the basket of goods and services. The main reason for their exclusion is the difficulty in determining their unit values and quantities. They are produced outside the market and there exist few goods and services in the marketplace for which comparable prices could be taken.

It was felt that, in the past, this has been an acceptable approach. However, as more and more goods and services are provided by governments, their share of the family's total consumption of goods and services further increases. A larger share of the family expenditure dollar is given for these goods either directly (such as by means of extra billing for health care services) or indirectly via the tax system. We were told by a number of persons that the impact on the cost of living of the financing of public goods is worth a thorough examination in the research program.

On the issue of taxes in general, it was pointed out many times that the inclusion of some taxes (such as retail sales taxes, exise taxes and property taxes) while purposely excluding other taxes (such as personal income taxes and payroll taxes) resulted in an obvious asymmetrical treatment of this very important component of the cost of living. If a government chose to generate additional revenue by an increase in the rate of personal income taxes, the CPI would not be affected. If on the other hand, it wished to effect a reduction in the CPI, it could reduce retail sales taxes or any other indirect tax for which it had jurisdiction. Statistics Canada was urged by some attendees to examine the matter as part of its review of the taxation/public goods issue in the research program.

(d) Special Indexes

There were frequent requests for additional indexes to be provided for special groups, particularly for **low income families and unattached individuals**. Even though an experimental index produced by the Statistics Canada Prices Division for the period 1978 to 1981 showed no effective difference over a period of years between the total CPI and the ''poor and aged CPI'', disbelief remained strong.

Counterarguments suggested that although this may have been the case in the past, it was thought that net income for these groups was no longer increasing at a rate comparable to the rate of change of the CPI and that expenditure patterns had changed significantly. A consumer price index reflecting more current expenditure patterns for these groups will diverge from the CPI, we were told.

Field pricing practices, as they pertain to the sample of stores selected for inclusion, were also thought to be a cause of important downward bias in the experimental indexes. Low income families and senior citizens supposedly are able to shop only in those stores readily accessible to the ''poorer parts of town'' where they are obliged to live because of their income level. These retail outlets are believed to have both a wider range of price movements and higher absolute prices as compared to their counterpart stores in medium and higher income areas. Thus, if pricing of goods for the construction of a consumer price index for low income families and senior citizens were restricted to such low income area retailers the resultant index would be consistently higher than the CPI.

Although knowledge of the Statistics Canada experimental Consumer Price Index and the recent record of its behaviour were not widespread, support for its continued production was frequently expressed.

Another group for which additional indexes were requested was the **non-urban population**. In each city visited there were representatives from provincial and municipal government social services programs who strongly believed that using either the national CPI, or the CPI for the closest metropolitan area, to index or otherwise make adjustments to transfer payments for non-urban residents is quite inappropriate. The national average was thought to benefit those who lived in areas with lower than average costs while residents

of areas where costs are higher do not gain from indexing to a national average. Residents of rural and remote areas face different living costs and, it was thought, face different price movements than do urban dwellers. Therefore, a large number of social assistance program administrators expressed their desire to see urban and non-urban consumer price indexes for their respective province or territory. There were requests for sub-provincial indexes to be produced more frequently and in sufficient detail to allow users to separate monthly figures into non-urban and urban indexes. This might, in their words, "reduce the indiscriminate use of the national CPI at all levels". As well, it might eliminate the common practice of "arbitrarily" adjusting the national level index to suit a variety of applications.

(e) Spatial Consumer Price Indexes

The question of measuring **relative price differences** between cities and more generally between regions came up frequently. The requests for such spatial indexes primarily came from persons employed by national firms who are involved in compensation policy matters on behalf of their company. As well, we were told that spatial indexes would be used at the sub-provincial level to adjust provincially administered transfer payments programs. However, urban/non-urban differentials seemed to be viewed as the more important consideration for sub-provincial income redistribution requirements.

It would appear that the private sector's interest in spatial indexes stems from two somewhat related applications. The first has to do with determining the amount of dollar compensation that would be sufficient to cover differences in living costs when relocating employees to other parts of Canada. The second application has to do with determining parity awards by arbitrators attempting to resolve wage or salary contract disputes. For example, does a professor at the University of British Columbia have to be paid "X" percent more than his counterpart at McGill because of the "higher cost of living" in Vancouver as compared to Montreal? What official data are available from Statistics Canada to put a value on "X"?

Often, contract negotiations are stalled because of a lack of credible information to help determine what constitutes equal pay in a **real sense as opposed to equal pay in a nominal** sense. A teacher earning $28,000 a year in Vancouver earns pay, equal in a **nominal** sense

to a teacher earning $28,000 a year in Fredericton. The important question is, however, do they earn the same in a **real** sense (that is, in the sense of effective buying power) given the apparent differences in the cost of living between the two cities? These are the kinds of questions for which better data are required to provide concrete answers. These are the questions people are having todeal with on a day-to-day basis.

Although the optimal solution would be for Statistics Canada to produce spatial cost-of-living indexes for cities (or regions), the participants requesting spatial indexes recognized the conceptual and operational difficulties in developing such measures. They recommended that a research effort be directed toward examining the problems of spatial indexes. In the meantime, or until such time as comprehensive spacial measures are produced, Statistics Canada should regularly make average price data available for cities that are currently surveyed in the existing CPI pricing system, we were told.

(f) The Cost of Consumer Credit

As we move toward a "cashless" society, the availability of credit to the consumer population will likely affect purchase patterns and even price sensitivity on the part of buyers. The increase in the use of credit for items such as food for consumption both inside and outside the home is causing significant shifts in the patterns of consumer expenditure. To fully understand the impact of this trend, we were told, Statistics Canada will need to examine the shifts in buying practices caused by the use of consumer finance services as well as the importance of consumer finance charges as a proportion of the expenditure dollar. If a significant and increasing portion of consumer expenditure is in fact being used for the purchase of the services of deferred payment from banks and other consumer finance institutions, many people were of the view that it may be inappropriate to continue to exclude such costs from the CPI.

To include the services of consumer finance in the CPI was generally viewed as an improvement since it was thought to result in a more cost-of-living-oriented Consumer Price Index. However, the implications of such a change did not pass unnoticed. Some participants acknowledged that to include consumer finance as a purchased consumer service may imply that the goods and services financed are investments. As such, an appropriate adjustment for a possible capital gain accruing to such "investments" must also be considered

as an element offsetting the cost of the "basket".

The issue of consumer credit was an important one for many of those participating in the discussions and they felt that it certainly warranted a re-examination in the proposed research program.

(g) Information Program

Statistics Canada's initiative in undertaking an open review of the measurement of aggregate price movement and to present independent research at a national public conference was fully endorsed. Our presence in the regions was welcomed as a genuine effort by Statistics Canada to involve the public in the defining of an appropriate research program.

Participants generally agreed that the CPI is often quoted but generally misunderstood. Further, they expressed a belief that Statistics Canada could take a pro-active stance toward reducing some of the misunderstandings about the CPI. At many meetings, questions concerning information access and information dissemination were presented. As well, there were constructive criticisms of the present system of presentation and interpretation of the monthly CPI data.

Immediate concern at the meetings focused on the need for a continuous information and education program at the "grass roots" or user level. Questions about the detailed underpinnings of the CPI are dealt with regularly by Statistics Canada on a one-to-one basis. However, it was felt that a system to make such information generally and readily available was needed - perhaps in a special publication or a series of articles.

Some of the kinds of topics the participants felt they would like to see addressed were:

- how to develop an index for a special group or for a special application (using Statistics Canada data);
- how to choose an index appropriate for a cost-of-living adjustment clause, along with suggested wordings for such clauses;
- the direct and indirect ramifications, if any, of the updated CPI for the existing labour contracts which refer to the CPI as at a particular date

the limitations of the CPI given its underlying concepts and definitions, and methodology.

Many users expressed a wish to construct indexes specific to their particular needs. In order to help users help themselves in the construction of special indexes, it was suggested that Statistics Canada make average price data available on a regular basis.

With regard to the monthly Consumer Price Index publication, it was suggested that the CPI information be presented alongside other related economic indicators. It was suggested that this might help place changes in the CPI in the context of the overall economic condition. (For example, a monthly index of per worker wages and salaries and an index of productivity might be presented along with changes in the CPI, in a graphic form.) Media reporting of the monthly results on a regional basis might be promoted if the regional highlights were provided in a format that would allow direct quotation and if tables and graphs were of a quality and size amenable to direct incorporation by the print media.

A request was made for the simultaneous availability of regional information on CPI release day. Under present practices the publication containing detailed regional data is not available at that time. There can be delays of up to one day in the receipt of detailed regional data.

(h) New Goods and Technological Change

One of the requirements of a Laspeyres index is to keep the ''weights'' associated with varying price regimes constant over time. For the CPI this means that the basket of goods priced and their corresponding contributions to the family expenditure pattern are kept constant and updated on the occasions when family expenditure data are available for a more current period. This brings to the fore the problems which occur when new items appear in the marketplace. How should the CPI be adjusted to reflect the new items?

It was expressed over and over that the CPI is being used as a cost-of-living index and that, given this use, the CPI should be a cost-of-living-oriented index. One major step in this direction we were told, would be to have the CPI reflect on a current basis the kinds of changes and substitutions that consumers make in their purchase patterns. Whatever the reasons for such changes, be they due to rising costs, changes in interests and lifestyles

or improvements in living standards, it is believed that they affect the cost of living and therefore should be included. Since the growing spectre of new goods, frequently made possible as a result of new technology, contributes substantially to changes in expenditure patterns, an examination of their treatment was recommended for the Statistics Canada research program.

In some instances, new goods are sold along with older versions of such goods. In other instances the new goods replace old goods completely in as much as the old goods are no longer produced. Then too there are new goods which are really **new**. In an effort to maintain a constant basket of goods, the treatment of the new goods can be different in each situation.

This leads to the question of the effects of quality change and the effects of rapidly advancing technology and why they constitute a matter for concern in the development of the CPI beyond their importance in the cost of living. As indicated earlier, the CPI is employed in many ways for which it was not designed. One of those ways is as a deflator in the measurement of real output and hence, of productivity. If for example, the effects of technical advancement were not fully captured in the specification and pricing of a good in the CPI basket, the resultant aggregate measure of price change, that is, the CPI, could be overstated. To then employ an overstated CPI as a deflator of nominal values of economic production would result in a consistent downward bias in the constant dollar measure of the national output.

All of this takes on an added emphasis at this time, in the minds of many persons, as a consequence of the widespread belief that the pace of technological change has quickened, especially as a consequence of the micro-processor based computer communications "revolution". The distinction between consumer goods and investment goods is being blurred (a home computer can be used to create financial models and to play various kinds of games). The distinctions between place of work and principal residence are also being blurred. What implications will all of this have for defining the consumer "basket"?

(i) Seasonal Variation

Representatives of those responsible for administering health and social services programs of the government of NWT indicated a potential problem with seasonal goods for consumer price indexes in the North. To some extent, the same problems could apply to the rest of Canada.

In areas isolated by large bodies of water (such as Yellowknife) transportation is regularly interrupted each spring for a six- to eight- week period when the icebridges break up and the water is too congested with ice to allow boat passage. During this period the prices of foodstuffs and other articles in short supply rise significantly. The inclusion of these ‘‘seasonally’’ higher prices in a consumer price index will create irregularities or disturbances in the series and may prompt requests for ‘‘seasonally adjusted’’ indexes. Other kinds of seasonal regularities occur in CPI series for other areas of Canada.

Since it appeared to be a widespread phenomenon, given the climatic conditions in different parts of Canada, the need for examining seasonality in price indexes was suggested as an area for research.

(j) Imported Inflation and International Comparability

Among the other areas of interest and concern were the issues of imported inflation and the somewhat related question of international standards for indexes of consumer prices.

In many discussions of the use of the CPI as a tool for indexing of incomes, the question of ‘‘imported’’ inflation was raised. Namely, should Canada promote the consumption of foreign goods within its boundaries by compensating its citizens for the purchase of imported goods for which prices are rising? The problem is particularly important in those instances where Canadian goods could be purchased more cheaply. If imported goods could be excluded from the CPI, it was suggested, the index would reflect only those price changes which occurred for Canadian goods and any use of the CPI as a tool for indexing incomes would be on the basis of ‘‘domestic’’ inflation.

A CPI is often regarded as a barometer of a country's economic well-being. The effects of fluctuations in consumer price indexes on a country's currency in the international money markets can be profound. Unfortunately the reactions of one country's financiers to changes in another country's CPI are based on the assumption that all CPIs are fundamentally the same and are therefore comparable.

For example, changes in the CPI in Canada are regularly compared to changes being reported for the United States. Differences in these rates of change are thought to reflect different states of economic health. However, variations could reflect differences in concepts and methods as much as differences in underlying price change. For this reason, Statistics Canada was encouraged to review concepts and methods in other countries, notably the USA, to ensure comparability and thus avoid spurious signals being given to the exchange markets.

IV. Conclusions

The objective of the user consultation process was to engage in bilateral discussions with users of price data:

- to inform them of the Price Measurement Review Program;
- to clarify any misunderstandings concerning the CPI;
- and to seek direction for a research program, specifically aimed at adding to our knowledge of the strengths and weaknesses of this important measure of economic performance.

These objectives were realized in full. The activity was completed in a relatively short period of time and on a modest budget.

After each meeting, participants commented that they were pleased to have had the opportunity to offer their comments and criticisms of the CPI. Furthermore, they indicated that they were leaving the sessions with an increased understanding of the CPI and they expressed an enthusiastic support for the Price Measurement Review Program.

The discussion of the various concerns for the CPI will provide an excellent source for direction in the shaping of the research program.

In our opinion, there is an important conclusion that should not be overlooked given the overwhelming support for and positive reaction to the consultation activity. Statistics Canada must continue the information and education process initiated in the consultative process.

Footnote

[1] Provincial quarterly indexes (excluding urban centres less than 30,000 population) are presently produced and distributed annually for all provinces. However, the distribution is limited to the Statistical Focal Points for internal use of the provincial statistics offices.

PROGRAMME D'EXAMEN DE LA MESURE DES PRIX: COMPTE RENDU DES RENCONTRES DE CONSULTATION

To provide you with a version in the official language of your choice, the French text is preceeded by the English text (p. 19) in this publication.

Cynthia Baumgarten
Programme de l'examen de la mesure des prix, Statistique Canada
Cyril D. Hodgins
Western Economic Service Ltd.

I. Introduction

Au cours des dernières années, des représentants de syndicats, d'associations de consommateurs, du commerce de détail et des gouvernements, ainsi que des économistes et des journalistes, ont fait part publiquement de leurs préoccupations quant à la validité de l'indice des prix à la consommation comme mesure principale de l'inflation au Canada. Il s'est également produit certains malentendus relatifs à la nature exacte de l'indice des prix à la consommation et à la réalité qu'il mesure. Dans le climat d'inflation que nous connaissons actuellement, on peut s'attendre à ce que des inquiétudes de ce genre continuent à se manifester, ce qui pourrait menacer l'utilité de l'indice comme mesure largement reconnue des mouvements mensuels des prix de détail.

Conscient de ces préoccupations, Statistique Canada a décidé de mettre sur pied un programme d'étude, d'évaluation et d'information visant les objectifs suivants:

a) établir un processus d'analyse en profondeur des diverses notions et méthodes qui servent à l'établissement de l'indice des prix à la consommation, analyses qui ou bien confirmeront que le processus actuel est le plus efficace, ou bien proposeront des mesures concrètes pour améliorer le calcul de l'indice, et

b) faire en sorte que les notions qui régissent la mesure des prix agrégatifs soient mieux connues du public, ce qui pourra contribuer à réduire la confusion chez les observateurs et le public en général quant à la nature de l'IPC et aux limites réalistes de ce qu'il peut et ne peut pas mesurer.

Le programme d'étude est constitué de trois volets:

1) Projets d'échange d'informations et de consultation des utilisateurs
2) Programme de recherche
3) Projet de conférence

Le Projet d'échange d'informations a consisté en une série de tables rondes grâce auxquelles des experts canadiens et étrangers ont pu se réunir et échanger leurs vues sur divers aspects de la mesure du niveau des prix et du processus inflationniste.

Le Projet de consultation des utilisateurs a consisté en une série de rencontres entre des représentants de Statistique Canada et un large éventail d'utilisateurs de l'IPC. Les rencontres avaient pour but d'examiner le programme de travail de Statistique Canada en matière de mesure des prix et de recevoir les commentaires, critiques et doléances des participants relativement aux mesures actuelles. Le programme de recherche, patronné par Statistique Canada, sera orienté en fonction des résultats de ces deux projets.

Le Projet de conférence consistera à organiser une conférence publique où seront présentés les résultats des recherches faites sur commande, ainsi que ceux d'autres travaux. Il est prévu de publier le texte des présentations, les commentaires des participants et un sommaire du compte rendu de la conférence.

L'organisation des rencontres de consultation avec les utilisateurs s'est faite avec l'aide des représentants de l'assistance-utilisateurs dans les bureaux régionaux et les bureaux secondaires de Statistique Canada, ainsi que des points de contact avec Statistique Canada des gouvernements des provinces et des territoires. On savait qu'en faisant appel à ces responsables, qui connaissent bien les principaux utilisateurs des indices de prix dans les régions et dans les divers gouvernements, on pourrait transmettre des invitations au plus grand nombre possible de personnes et d'organismes susceptibles d'être intéressés au programme. Le choix des intervenants à rencontrer a été, en définitive, un processus arbitraire. En fait, seule la tenue d'audiences publiques aurait permis d'éviter le caractère arbitraire du choix des intervenants, mais on ne disposait ni du temps ni des fonds nécessaires pour organiser de telles audiences. Il n'en reste pas moins que la méthode adoptée a contribué à entamer, avec les utilisateurs intéressés, un dialogue sur une partie très importante de la production de données statistiques.

Les ministères et agences du gouvernement fédéral qui désiraient participer activement au processus de consultation ont été invités à envoyer des représentants à une réunion concernant ce sujet.

Ce rapport présente un sommaire des principales conclusions qui sont ressorties des rencontres de consultation des utilisateurs.

II. Rencontres de consultation: contenu et participation

Des rencontres ont eu lieu dans les dix provinces, ainsi qu'au Yukon et dans les Territoires du Nord-Ouest. En tout et pour tout, on a tenu 29 rencontres dans 15 villes au cours de la période du 25 janvier 1982 au 14 mai 1982.

Le nombre total de participants, par secteur ou affiliation, est indiqué au tableau 1. Plus de la moitié des personnes consultées étaient des fonctionnaires de ministères ou d'organismes provinciaux. Vient ensuite le secteur des affaires, suivi du gouvernement fédéral et des syndicats ouvriers. Dans la catégorie ''autres participants'', on trouve des groupes comme l'Association des consommateurs du Canada. Malgré les efforts de Statistique Canada, très peu de représentants des médias se sont présentés. En fait, ceux-ci ne constituaient qu'une petite fraction de ceux qui avaient été avisés des rencontres et invités à y participer.

Les représentants de Statistique Canada commençaient chaque rencontre par un court exposé, sans formalités. D'une rencontre à l'autre, on a varié légèrement cette présentation pour insister sur des points différents, surtout pour mettre en lumière des problèmes ou des préoccupations locales, mais on a développé principalement les idées suivantes.

L'inflation est la principale cible de la politique monétaire et fiscale, tant au Canada qu'aux États-Unis, et le principal sujet de préoccupation des marchés du crédit en Amérique du Nord. C'est également au nom de l'inflation que plusieurs craignent tant les déficits budgétaires des gouvernements canadien et américain, à cause des conséquences inflationnistes que ces déficits peuvent avoir. Mais qu'est-ce donc que l'inflation?

Selon les définitions qu'on peut trouver dans les manuels, il s'agit d'une hausse soutenue du niveau général ou global des prix. Elle est soutenue en ce sens qu'elle est continue, c'est-à-dire qu'il ne s'agit pas d'une variation occasionnelle ou périodique. Elle s'applique au niveau général des prix, par opposition aux mouvements de prix relatifs qui résultent de changements des conditions de l'offre et de la demande de biens ou de services particuliers comme les oranges, les chaussures pour dames ou l'essence.

Le taux d'inflation se mesure de bien des façons. On dispose d'indices sur le mouvement du niveau moyen des prix demandés par l'industrie et d'indices sommaires sur les prix accordés aux agriculteurs. Il existe des indices de prix implicites qui sont calculés à partir des valeurs en dollars courants et en dollars constants des composantes de la dépense nationale brute.

Et il y a bien entendu l'indice des prix à la consommation, celui qui dans l'esprit de la plupart des Canadiens est l'image du taux d'inflation. Il semble que l'IPC ait mérité le titre de mesure définitive du taux d'inflation pour plusieurs raisons:

- L'IPC décrit l'évolution générale des prix des biens et des services que nous, les consommateurs, nous procurons quotidiennement sur le marché.
- L'IPC est publié chaque mois avec peu de retard.
- L'IPC peut être décomposé en plusieurs éléments, de telle sorte que les mouvements d'ensemble peuvent être expliqués par les variations de composantes particulières.
- L'IPC est calculé pour quelques principaux centres urbains du Canada.

Les autres mesures de l'inflation, comme les indices implicites de déflation de la DNB, sont loin d'être aussi utilisées et analysées. Ce genre d'indice ne semble pas connu du public en général et n'est publié qu'à tous les trimestres. Les indices implicites de déflation des comptes nationaux des revenus et des dépenses ne peuvent être décomposés dans le détail et ne sont calculés qu'à l'échelle nationale. L'IPC est donc devenu le critère d'évaluation du taux d'inflation dans l'esprit de la plupart des Canadiens. Bien plus, on s'y reporte largement pour évaluer l'évolution du coût de la vie, et son utilisation s'est généralisée dans plusieurs contextes en vue de la protection du pouvoir d'achat réel ou effectif de toute une variété de paiements. Cela s'est manifesté sous la forme d'indexation des salaires dans le sconventions collectives, ou encore sous forme d'indexation des paiements de transfert

des gouvernements comme les allocations familiales ou les prestations de sécurité de la vieillesse. On s'est également servi de l'indice pour instaurer des ajustements automatiques de la structure d'imposition du revenu des particuliers, pour éviter que les gouvernements ne profitent des retombées de l'application de taux d'imposition progressifs à des revenus gonflés par l'inflation.

À cause de son importance, l'IPC est certes le plus en vue de tous les indicateurs de la performance économique produits par Statistique Canada. Les bureaux régionaux de Statistique Canada reçoivent plus de demandes de renseignements sur l'IPC que sur toute autre mesure statistique particulière.

Il ne fait aucun doute que l'importance de l'IPC, ainsi que la variété des usages auxquels il se prête, vont bien au-delà de ce qu'auraient pu imaginer ceux qui l'ont élaboré au départ. C'est en 1913 qu'on a commencé au Canada à calculer l'IPC, dont l'appellation officielle a été, jusqu'en 1949, indice du coût de la vie.

La plupart des utilisateurs de l'indice connaissent sa définition, soit la variation dans le temps, exprimée en pourcentage, du coût d'un panier constant de biens et de services, qui représente les achats faits par un groupe particulier de la population au cours d'une période donnée. Comme le panier doit renfermer des biens et des services de quantité et de qualité invariables ou comparables et dont les variations de prix sont mesurables dans le temps, les fluctuations de son coût devraient résulter uniquement du mouvement des prix.

Bien que cette description particulière porte le sceau de la neutralité professionnelle, la mesure telle qu'elle est établie actuellement n'est pas exempte de critiques. Un bon nombre d'inquiétudes formulées à son endroit sont bien connues et traduisent, dans bien des cas, le souhait que l'indice soit davantage un indice du coût de la vie qu'il ne l'est actuellement. D'aucuns prétendent que l'IPC surestime le taux de variation du coût de la vie parce qu'il ne tient pas compte des substitutions effectuées par les consommateurs qui achètent moins les produits dont le prix relatif a augmenté et optent pour ceux qui sont offerts en réduction, à des prix spéciaux ou à des prix de liquidation. D'autres par contre prétendent que l'IPC **sous-estime** le taux de variation du coût de la vie parce qu'il accorde trop peu d'importance, dans le panier, à des composantes comme l'énergie et le logement qui ont connu récemment une escalade des prix.

Bien que certaines personnes puissent penser que l'IPC prête à confusion ou qu'il faut carrément l'écarter, il n'est probablement pas faux d'affirmer que la majorité des Canadiens acceptent l'information qu'il nous donne quant à l'évolution des prix, car nous ne disposons pas véritablement d'une solution de rechange. Il n'existe aucune autre mesure avec laquelle l'indice puisse être directement comparé.

Depuis des années, Statistique Canada applique son propre programme continu de recherche et de développement de manière à s'assurer que les notions et les méthodes sont adaptées aux circonstances nouvelles. Toutefois, compte tenu de l'importance considérable que revêt actuellement l'IPC, le statisticien en chef a décidé d'ajouter cette année au programme de recherche interne de Statistique Canada un nouveau programme d'étude de la mesure du niveau des prix. Ce programme comporte 3 volets principaux:

- une série de consultations bilatérales avec les spécialistes et les principaux utilisateurs de toutes les régions du Canada;
- l'appui de Statistique Canada à un ensemble d'études spéciales devant être effectuées par des économistes professionnels réputés des milieux universitaires; et
- une conférence nationale sur la mesure du niveau des prix.

Cette démarche globale traduit un esprit d'ouverture face à la possibilité d'améliorer le processus de mesure par suite d'un échange public d'opinions professionnelles sur des sujets les plus complexes ayant trait aux notions et aux méthodes. Elle n'a aucunement été entreprise parce qu'on aurait découvert récemment une anomalie quelconque dans l'établissement de l'IPC.

Il ressort des commentaires qui nous ont été formulés au cours des rencontres de consultation avec les utilisateurs que cette initiative de Statistique Canada a été très bien reçue et largement appuyée partout au pays. Un nombre important d'économistes réputés des milieux universitaires ont fait savoir qu'ils étaient prêts à participer par le biais de travaux de recherche.

Plusieurs groupes se sont révélés disposés à faire valoir leurs vues, en exprimant les préoccupations qu'ils pouvaient avoir au sujet de l'IPC et de son mode de calcul, et en se disant

intéressés à prendre part à la conférence. Nous avons confiance, grâce à cet intérêt et à cet appui, de faire de ce programme une réussite.

Il est à espérer que le succès obtenu incitera Statistique Canada à entreprendre, dans d'autres secteurs statistiques d'importance, des programmes semblables de soutien à la recherche menant à des conférences publiques. Une telle initiative d'un organisme fédéral pourra, il faut l'espérer, paver la voie à d'autres démonstrations d'ouverture et de responsabilité à l'égard du public, de la part d'autres secteurs du gouvernement fédéral.

On trouvera dans les sections qui suivent un résumé de certains des commentaires formulés par les utilisateurs.

Tableau 1. Participation aux rencontres de consultation des utilisateurs du programme de révision de la mesure des prix

Secteurs d'activité	Nombre de participants	Pourcentage
Syndicats ouvriers	16	6
Secteur des affaires	46	17
Médias	3	1
Gouvernement fédéral*	39	15
Gouvernements provinciaux		
T.-N.	6	2
N.-É.	4	1
N.-B.	13	5
Î.-P.-É.	13	5
Qué.	18	7
Ont.	15	6
Man.	10	4
Sask.	7	3
Alb.	16	6
C.-B.	14	5
Yukon	11	4
T.N.-O.	13	5
Total – Gouvernements provinciaux	140	53
Autres participants	19	7
Total des participants	263	100

* Sans les représentants de Statistique Canada (par ex. des bureaux régionaux).

III. Mesure du niveau des prix

La présente section résume les principaux sujets de préoccupation dont ont fait part les utilisateurs au cours du processus de consultation. L'ordre de présentation ne correspond aucunement à un ordre de priorité ou d'urgence. En fait, la priorité accordée aux différents problèmes semblait varier considérablement selon les intérêts particuliers des participants et la région du Canada où avait lieu la rencontre.

a) Fréquence des mises à jour

En avril 1982, un schéma de pondération tenant compte des habitudes de dépense des consommateurs de 1978 a remplacé celui de 1974 qui avait servi jusqu'alors de base pour déterminer l'importance relative, dans le panier de l'IPC, des produits achetés par les consommateurs. Les participants aux rencontres ont exprimé l'avis que le fait de se référer à 1978 plutôt qu'à 1974 ne constituait qu'une amélioration partielle du processus de pondération. À maintes reprises, ils se sont montrés préoccupés de ce que les poids n'étaient pas suffisamment "actuels". À titre d'exemple, les consommateurs ont vu grimper, depuis 1978, les coûts de l'énergie et des produits dont la fabrication consomme beaucoup d'énergie. Ils ont donc modifié en conséquence leurs habitudes de dépense mais, nous a-t-on signalé, l'effet de ce changement n'apparaît pas dans l'IPC actuel et n'apparaîtra pas non plus dans l'IPC après la mise à jour à venir. De l'avis de tous, des problèmes de ce genre minent la crédibilité de l'indice.

Une version plus actuelle des habitudes de dépense, même s'il ne s'agissait que d'une amélioration rudimentaire, valait quand même la peine d'être examinée. La plupart des participants avaient l'impression que la crédibilité de l'indice s'en trouverait augmentée si l'on établissait un processus non arbitraire de mise à jour des schémas de pondération lorsque la répartition des éléments qui composent le budget familial semble s'être modifiée de façon significative. On a souvent proposé comme solution la possibilité d'établir des diagrammes de pondérations estimatives, fondés sur d'autres indicateurs principaux pertinents, récents et disponibles.

On considère que les enquêtes sur les dépenses des familles, qui sont un des éléments fondamentaux servant à l'établissement de l'IPC, peuvent être un moyen possible de rendre plus actuelle la pondération sur laquelle est fondé l'IPC. L'idée d'un examen global des notions et des méthodes sous-jacentes à ces enquêtes, à la lumière des besoins de l'IPC, a été accueillie favorablement.

b) Coûts du logement

Dans chaque ville, les utilisateurs ont soulevé des problèmes concernant la composante logement de l'IPC. On a exprimé l'avis que les sous-indices du logement en propriété et en location, et particulièrement les sous-composantes de l'indice des loyers, n'exprimaient pas la réalité du marché. En fait, on a dit croire que les augmentations des prix du logement en propriété et en location sont beaucoup plus fortes que celles indiquées par les sous-indices correspondants de l'IPC. L'indice des prix du logement du Trust Royal et les données sur les loyers de la SCHL viennent confirmer cette impression. On nous a signalé que ces deux sources montrent des augmentations supérieures à celles qu'indiquent les composantes comparables de l'IPC, au niveau des villes.

La méthode actuelle qui consiste à appliquer une distribution fixe des hypothèques en cours selon l'âge sur une période de 60 mois donne, selon les avis formulés, une vue inférieure à la réalité de la progression des coûts des utilisateurs et entraîne un retard dans la prise en compte des fluctuations des taux d'intérêt hypothécaires. Depuis quelque temps, il se négocie sur le marché des contrats hypothécaires de beaucoup plus courte durée et certaines banques offrent maintenant des taux d'intérêt hypothécaires qui fluctuent d'un mois à l'autre. Les acheteurs de logements neufs se voient souvent offrir des taux d'intérêt hypothécaires réduits (inférieurs aux taux du marché) pour des maisons dont le prix de vente comporte une marge couvrant la réduction de l'hypothèque. Dans quelle mesure ces nouvelles pratiques sont-elles répandues et quel serait leur effet sur le taux de variations de la composante logement en propriété de l'IPC si les méthodes étaient modifiées pour tenir compte de ces pratiques, voilà des questions qui ont été soumises à titre d'objets possibles de recherche.

Toujours en ce qui concerne le logement, une autre question portait sur la formule actuelle de calcul des coûts d'un logement en propriété. On a fait valoir que la pratique actuelle de ne pas pondérer les gains en capital en fonction du coût de la maison surestime la variation du "coût réel" que les propriétaires doivent supporter.

c) Traitement des impôts, taxes et biens publics

Le montant de l'impôt sur le revenu payé n'étant relié à aucune quantité particulière de biens ou de services reçus par la population, ceux-ci ne sont pas inclus dans le calcul de l'IPC. Selon de nombreux participants, bien qu'il soit juste dans certains cas particuliers de procéder ainsi pour définir un indice des prix à la consommation, on devrait quand même en tenir compte dans l'IPC étant donné que ce dernier est utilisé comme indice du coût de la vie. Il a été souligné que l'impôt sur le revenu accapare une portion substantielle du revenu d'une famille et représente donc un facteur très important du coût de la vie. Comme on se sert très largement de l'IPC pour indexer les salaires et les paiements de transfert, il faudrait d'après ce raisonnement, construire un indice qui soit un reflet plus fidèle du coût de la vie.

Certains ont proposé que Statistique Canada élabore un système d'indices dont certains, une fois amalgamés, pourraient former un indice approximatif du coût de la vie. L'IPC dans sa forme actuelle serait une composante de ce système.

Les "biens publics" (c'est-à-dire les biens et services fournis à la population par l'État) sont payés par le biais du régime fiscal. Bien que les recettes fiscales puissent servir en entier à la production de biens publics, ces derniers ne sont pas tous des biens de consommation. On pourrait donc prétendre que l'IPC devrait tenir compte du coût des biens publics consommés plutôt que du coût des biens publics produits. Autrement dit, l'inclusion des impôts et taxes dans le calcul d'un indice dont le but est d'exprimer le coût de la vie risque de surestimer (ou de donner un poids trop élevé) à la composante "biens publics".

Si on ne tient absolument pas compte de biens publics dans le panier de biens et de services qui sert au calcul de l'IPC, c'est surtout parce qu'il est difficile de déterminer leurs valeurs et leur quantité. Ils sont produits à l'extérieur du marché et il existe peu de biens et de services comparables, sur le marché, dont les prix puissent être observés.

Cette exclusion a paru jusqu'ici acceptable, mais comme une quantité toujours plus considérable de biens et de services est fournie par les gouvernements, leur part s'accroît dans la consommation totale des familles. Une plus grande fraction de chaque dollar dépensé est consacrée à ces biens, soit directement (par exemple sous forme de frais supplémentaires à payer pour des services de santé), soit indirectement par l'intermédiaire du système fiscal. Plusieurs ont exprimé l'avis qu'il valait la peine, dans le cadre du programme de recherche, d'examiner en profondeur l'impact du financement des biens publics sur le coût de la vie.

En ce qui concerne l'ensemble des impôts et taxes, il a été mentionné à maintes reprises que le fait d'inclure certaines taxes (par ex. taxe de vente au détail, impôt foncier) et d'en exclure d'autres (par ex. impôt sur le revenu des particuliers) constituait manifestement un traitement asymétrique de cette composante très importante du coût de la vie. Si un gouvernement décidait de tirer des recettes supplémentaires en augmentant les taux d'imposition des particuliers, l'IPC ne serait pas touché. Si, par contre, il désirait faire baisser l'IPC, il pourrait diminuer les taxes de vente au détail, ou toute autre taxe relevant de sa compétence. Il a été fortement recommandé que Statistique Canada étudie cette question dans le volet de son programme de recherche traitant des impôts, taxes et biens publics.

d) Indices spéciaux

On a souvent demandé que soient calculés des indices supplémentaires visant des groupes spéciaux, notamment les familles à faible revenu et les personnes seules. Bien que l'indice expérimental produit par la Division des prix pour la période de 1978 à 1980 n'ait démontré aucune véritable différence sur plusieurs années entre l'IPC d'ensemble et "l'IPC des pauvres et des personnes âgées", nombreux sont ceux qui demeurent sceptiques.

On invoque comme argument que le revenu net de ces groupes, contrairement à la situation passée, n'augmente plus à un rythme comparable à celui de l'IPC et qu'il en résultera une modification sensible des habitudes de dépense, de sorte qu'un indice des prix à la consommation fondé sur des données plus actuelles quant aux habitudes de dépense de ces groupes s'écartera de l'IPC.

On a également exprimé l'avis que les relevés de prix sur le terrain, c'est-à-dire dans les magasins choisis pour faire partie de l'échantillon, étaient la source d'un important biais par défaut dans les indices expérimentaux. Les familles à faible revenu et les personnes âgées font, semble-t-il, tous leurs achats dans les magasins des quartiers "pauvres" où ils habitent parce qu'ils n'ont pas les moyens de vivre ailleurs. Les points de vente en question accusent des mouvements de prix plus amples et affichent des prix absolus plus élevés que les magasins équivalents des secteurs des classes moyenne et aisée. En conséquence, si le champ d'observation servant à l'établissement d'un indice des prix à la consommation pour les familles à faible revenu et les personnes âgées se limitait aux points de vente desservant une clientèle à faible revenu, il en résulterait un indice régulièrement plus élevé que l'IPC.

Par ailleurs, même si l'indice expérimental et les résultats qu'il a permis d'obtenir n'étaient pas très connus, on a fréquemment exprimé le souhait que cet indice continue d'être établi.

Il y a un autre groupe pour lequel des indices supplémentaires ont été réclamés, celui de la population non urbaine. Dans toutes les villes visitées, il y avait des représentants des programmes de services sociaux des administrations provinciales et municipales qui croyaient fermement qu'il ne convient pas d'utiliser soit l'IPC national, soit l'IPC de la région métropolitaine la plus proche, pour indexer ou ajuster de quelque autre façon les paiements de transfert visant la population non urbaine. Selon eux, la moyenne nationale profite aux résidents des régions où les coûts moyens sont plus faibles, mais elle désavantage ceux des régions où les coûts sont plus élevés. Dans les régions rurales et les régions éloignées, les coûts sont différents et, prétend-on, les mouvements des prix ne sont pas les mêmes que ceux des régions urbaines. Par conséquent, de nombreux administrateurs de programmes d'assistance sociale ont dit souhaiter que soient établis des indices urbains et non urbains des prix à la consommation pour leur province ou territoire. On a demandé que des indices intraprovinciaux[1] soient produits plus fréquemment et de façon suffisamment détaillée pour que les utilisateurs puissent répartir les données mensuelles en indices urbains et non urbains. On pourrait ainsi, selon les propres termes des intervenants, "réduire l'utilisation sans discernement de l'IPC national à tous les niveaux". Cela permettrait également d'éliminer la pratique courante qui consiste à ajuster de façon arbitraire l'indice national en fonction de toute une gamme d'applications.

e) Indices spatiaux des prix à la consommation

On a souvent fait allusion au problème de la mesure des différences de prix entre les villes, et plus généralement entre les régions. Ce sont surtout les personnes participant à l'établissement de politiques de salaires et de traitements, pour le compte des entreprises nationales qui les emploient, qui on réclamé de tels indices spatiaux. Ceux-ci pourraient également être utilisés, nous a-t-on dit, au niveau intraprovincial pour l'ajustement des programmes de paiements de transfert administrés par les provinces. On semblait toutefois considérer que ce qui importait le plus, à l'échelle d'une province, était de bien différencier les régions urbaines et les régions non urbaines.

Il semble que l'intérêt du secteur privé pour des indices spatiaux résulte de deux besoins qui sont liés d'une certaine façon. Le premier est celui de déterminer la rémunération différentielle, en fonction du coût de la vie, qu'il convient de verser à des employés mutés. Le deuxième a trait aux décisions sur la parité salariale que doivent prendre les arbitres dans les conflits contractuels.

Il arrive souvent que des négociations collectives achoppent parce qu'on ne peut se baser sur aucune information valable pour déterminer ce qui constitue un salaire égal au sens réel par opposition à un salaire égal au sens nominal. Des enseignants qui gagnent 20 000 $ par année ont le même salaire, au sens nominal, qu'ils travaillent à Vancouver ou à Fredericton. Mais qu'en est-il de leur salaire réel (c'est-à-dire de leur pouvoir d'achat réel) compte tenu des différences apparentes du coût de la vie entre les deux villes? C'est pour répondre concrètement aux questions de ce genre, qui se posent quotidiennement, qu'il faut produire de meilleures données, nous a-t-on dit.

Bien que la solution idéale serait que Statistique Canada produise des indices spatiaux du coût de la vie pour les villes (ou les régions), les participants qui les ont réclamés ont reconnu les difficultés théoriques et pratiques de l'établissement de tels indices. Ils ont recommandé avec insistance que soient effectués des travaux de recherche sur les problèmes que posent les indices spatiaux et que Statistique Canada fournisse, jusqu'à ce que de tels indices soient produits, des données régulières sur les prix moyens dans les villes qui font l'objet d'observations en vertu de la méthode actuelle de calcul de l'IPC.

f) Le coût du crédit à la consommation

Nous vivons de plus en plus à l'époque du ''crédit'', ce qui ne peut qu'influer sur les habitudes d'achat des consommateurs et même sur la sensibilité des acheteurs à l'égard des prix. L'utilisation croissante du crédit pour des articles comme les aliments consommés à la maison et à l'extérieur entraîne des modifications sensibles des habitudes de dépenses des consommateurs. Pour bien saisir la portée de cette tendance, Statistique Canada devra examiner les changements associés à l'utilisation de services de crédit à la consommation, ainsi que la part des coûts du crédit à la consommation dans chaque dollar dépensé. S'il s'avère que les consommateurs consacrent une part substantielle et croissante de leurs dépenses à l'achat de services de paiement différé auprès des banques et des autres institutions de prêt, il serait peut-être utile, de l'avis de plusieurs, d'inclure ces services dans l'IPC.

On considère en général que l'inclusion de ces services serait une amélioration, car elle contribuerait à produire un indice des prix à la consommation davantage axé sur le coût de la vie. Toutefois, les conséquences d'un tel changement n'ont pas été passées sous silence. Certains participants ont reconnu qu'en incluant le crédit à titre de produit de consommation, on considère que les biens et les services financés sont des investissements. Il faudrait donc prévoir un traitement approprié des éventuels gains en capital découlant de tels investissements afin de déduire ces gains du coût du ''panier''.

Il a également été souligné que le moment et la fréquence des mises à jour de la pondération selon les dépenses, ainsi que la fréquence d'observation des prix du service, pourraient avoir un effet sensible sur l'indice en périodes d'inflation rapide.

En bref, le crédit à la consommation est, aux yeux d'un grand nombre de participants, un facteur important qui mérite certainement d'être réexaminé dans le cadre du programme de recherche proposé.

g) Programme d'information

Tous les participants ont appuyé l'initiative de Statistique Canada de procéder publiquement à une étude de la mesure du mouvement des prix agrégatifs et de présenter les résultats de recherches indépendantes à une conférence publique nationale. Notre présence

dans les régions a été appréciée et perçue comme un effort authentique de la part de Statistique Canada de faire participer le public à la définition d'un programme de recherche approprié.

De l'avis général des participants, l'IPC est souvent cité, mais plutôt méconnu. Selon eux, Statistique Canada pourrait contribuer activement à éclaircir certains des malentendus qui existent au sujet de l'indice. À maintes reprises, des questions ont été posées relativement à l'accès à l'information et à la diffusion de cette dernière. On a également formulé des critiques constructives quant au mode actuel de présentation et d'interprétation des données mensuelles de l'IPC.

Au nombre des préoccupations immédiates, on a surtout fait valoir le besoin d'un programme continu d'information et d'initiation au niveau de ''la base'', c'est-à-dire des utilisateurs. Les divers éléments constituant le fondement de l'IPC sont traités régulièrement, un à un, par Statistique Canada. On a toutefois exprimé l'avis que la diffusion de cette information, pour être universelle et facilement accessible, devrait se faire de façon structurée, peut-être sous forme d'une publication ou d'une série d'articles.

Parmi les sujets que les participants aimeraient voir traiter, mentionnons les suivants:

- méthode d'établissement d'un indice pour une application ou un groupe spécial (à partir des données de Statistique Canada);
- méthode de sélection d'un indice approprié pouvant servir dans des clauses d'ajustement en fonction du coût de la vie, avec suggestions de libeller de telles clauses;
- éventuelles conséquences directes et indirectes d'une mise à jour de l'IPC pour les conventions collectives en vigueur qui font référence à l'IPC à une date particulière; et,
- limites de l'IPC compte tenu des concepts, définitions et méthodes sur lesquels il repose.

De nombreux utilisateurs ont exprimé le souhait de pouvoir construire des indices en fonction de leurs besoins particuliers. Il a été suggéré, afin d'aider ces derniers à produire des indices spéciaux, que Statistique Canada fournisse de façon régulière des données sur les prix moyens.

En ce qui a trait à la publication mensuelle de l'indice des prix à la consommation, on a proposé que les valeurs de l'IPC soient présentées vis-à-vis d'autres indicateurs économiques connexes, de façon que l'IPC puisse être associé à la situation économique globale. (On pourrait par exemple présenter avec les mouvements de l'IPC, sous forme de graphique, un indice mensuel des salaires et traitements par travailleur et un indice de la productivité.) La diffusion par les médias des données régionales mensuelles pourrait être facilitée si les faits saillants des régions étaient fournis dans une forme permettant de les citer directement et si les tableaux et graphiques étaient d'une qualité et d'une grandeur se prêtant à l'incorporation directe dans les médias imprimés.

On a demandé que les données régionales soient diffusées simultanément le jour de la publication. À l'heure actuelle, le document renfermant les données régionales détaillées ne paraît pas en même temps. Ces données régionales détaillées parviennent parfois à destination avec des retards d'une journée.

h) Produits nouveaux et changements technologiques

Une des exigences d'un indice de Laspeyres est que les "poids" associés à des régimes de prix variables demeurent constants dans le temps. Dans le cas de l'IPC, cela signifie que le panier des biens observés et leur part respective dans les dépenses des familles sont maintenus constants et mis à jour lorsqu'on dispose de données plus récentes sur les dépenses des familles. Mais qu'arrive-t-il des articles nouveaux offerts sur le marché? Comment l'IPC doit-il réagir pour en tenir compte?

Les participants n'ont pas cessé de faire valoir que l'IPC est utilisé comme indice du coût de la vie et qu'en conséquence, il devrait être axé davantage sur la mesure du coût de la vie. Au dire des participants, un pas important serait franchi dans cette direction si on faisait en sorte que l'IPC incorpore les données courantes sur les changements et les substitutions effectués par les consommateurs dans leurs habitudes d'achat. On croit que ces changements influent sur le coût de la vie et devraient donc être inclus, peu importe les raisons qui les ont motivés, qu'il s'agisse de la hausse des coûts, d'une évolution des goûts et des modes de vie ou d'une amélioration du niveau de vie. Étant donné que l'éventail toujours plus grand des produits nouveaux, qui sont souvent le résultat de l'application de technologies nouvelles, contribue sensiblement à la modification des habitudes

de dépenses, il a été recommandé que Statistique Canada examine le traitement qui leur est réservé dans le cadre de son programme de recherche.

Il y a certains nouveaux produits qui viennent s'ajouter à des produits analogues déjà existants. Dans d'autres cas, des nouveaux produits viennent remplacer ceux qui sont disparus du marché. Enfin, il y a des produits qui sont réellement des nouveautés. Afin de maintenir le panier constant, le traitement des nouveaux produits peut être différent dans chaque situation.

Cela nous amène à nous interroger sur les effets des variations de qualité et sur les conséquences d'une technologie qui évolue rapidement, et à nous demander pourquoi ils constituent un sujet de préoccupation pour l'établissement de l'IPC, par-delà leur importance en ce qui touche le coût de la vie. Comme on l'a dit plus haut, l'IPC est utilisé à plusieurs fins pour lesquelles il n'a pas été conçu. On l'utilise notamment comme indice de déflation pour mesurer la production réelle, donc la productivité. Si, par exemple, la spécification et le relevé de prix d'un produit du panier visé par l'IPC ne tenaient pas parfaitement compte des progrès technologiques, la mesure agrégative du mouvement de prix, c'est-à-dire l'IPC, serait surestimée et son emploi comme indice de déflation des valeurs nominales de la production économique entraînerait un biais par défaut continuel de la mesure de la production nationale en dollars constants.

Toutes ces questions semblent prendre aujourd'hui une importance accrue dans l'esprit de nombreuses personnes, dans le contexte d'une évolution technologique qui, de l'avis général, va en s'accélérant, par suite surtout de la ''révolution'' des communications informatiques engendrée par les microprocesseurs. La distinction entre biens de consommation et biens d'investissement s'estompe (un ordinateur domestique peut servir à créer des modèles financiers et à jouer à différents jeux électroniques). On fait également de moins en moins la distinction entre lieu de travail et résidence principale. Quels seront les effets de tous ces changements sur la définition du ''panier du consommateur''?

i) Variations saisonnières

Les représentants des administrateurs des programmes de santé et de services sociaux

du gouvernement des T.N.-O. ont fait part d'un problème susceptible d'affecter les indices des prix à la consommation des régions du Nord, soit celui des biens saisonniers. Le même problème peut se poser dans une certaine mesure dans le reste du Canada.

Dans les régions isolées par de grandes étendues d'eau (par ex. Yellowknife), le transport est régulièrement interrompu pendant une période de six à huit semaines au moment de la déblâcle du printemps, lorsque la congestion des glaces est trop forte pour permettre le passage des bateaux. Pendant cette période, les prix des aliments et d'autres articles dont les provisions sont peu abondantes connaissent une hausse sensible. L'inclusion aux fins du calcul de l'indice des prix à la consommation de ces valeurs ''saisonnières'' plus élevées provoquera des anomalies et des perturbations dans les séries, d'où il résultera éventuellement une demande d'indices ''désaisonnalisés''. Des disparités saisonnières de nature différente affectent les séries de l'IPC dans d'autres régions du Canada.

Comme la préoccupation semble répandue, ce qui n'a rien d'étonnant compte tenu des conditions climatiques des différentes régions du Canada, il a été suggéré que l'examen du facteur saisonnier dans les indices de prix soit inclus dans le programme de recherche.

j) Inflation importée et comparaison avec les autres pays

Parmi les autres sujets d'intérêt soulevés, il y a le problème de l'inflation importée et la question connexe des normes internationales relatives aux indices des prix à la consommation.

Il est arrivé à maintes reprises, dans les discussions sur l'IPC en tant qu'outil d'indexation des revenus, que la question de l'inflation ''importée'' soit soulevée. On se demandait en fait si le Canada devait encourager la consommation de biens étrangers au pays en indemnisant ses citoyens lorsqu'ils achètent des biens importés dont les prix montent. Le problème se pose de façon plus aiguë dans les cas où il existe des biens canadiens qui sont moins chers. Si on excluait les biens importés du calcul de l'IPC, celui-ci ne témoignerait semble-t-il que des mouvements de prix des biens canadiens; son utilisation pour l'indexation des revenus ne serait alors basée que sur notre propre inflation intérieure.

L'IPC est souvent considéré comme un baromètre de la santé économique d'un pays. L'effet des fluctuations des indices des prix à la consommation sur une monnaie nationale peut être considérable sur les marchés monétaires internationaux. Malheureusement, les experts financiers d'un pays réagissent aux variations de l'IPC d'un autre pays en supposant que tous les IPC sont fondamentalement identiques et donc comparables.

On compare constamment, par exemple, l'évolution de l'IPC au Canada à celle de l'IPC aux États-Unis. Les différences observées sont considérées comme révélatrices des différents états de santé de l'économie. Toutefois, il est possible que les variations soient le reflet de différences au niveau des concepts et des méthodes tout aussi bien que de différences dans l'évolution des prix comme tel. C'est pourquoi on a incité Statistique Canada à examiner les concepts et méthodes utilisés dans d'autres pays, notamment aux États-Unis, pour s'assurer que les indices sont comparables et ainsi éviter que les cambistes ne se basent sur de fausses indications.

IV. Conclusions

Le processus de consultation des utilisateurs devait permettre d'entamer des discussions bilatérales avec les utilisateurs des données sur les prix afin:

- de leur transmettre des informations sur le Programme de révision de la mesure des prix;
- de clarifier tout malentendu concernant l'indice des prix à la consommation;
- de tenter de donner une orientation à un programme de recherche destiné avant tout à mieux nous faire connaître les forces et les faiblesses de cette importante mesure de la performance économique.

Tous les objectifs de la consultation ont été pleinement atteints dans un temps relativement court et dans les limites d'un modeste budget.

Après chaque recontre, les participants se sont dit satisfaits d'avoir eu l'occasion d'exprimer leurs commentaires et leurs critiques relativement à l'IPC. Ils ont en outre signalé qu'à la suite des séances ils se sentaient mieux informés au sujet de l'IPC et ils ont accordé un appui enthousiaste au Programme d'étude de la mesure des prix.

Les discussions qui ont eu lieu sur les divers sujets qui préoccupaient les intervenants constitueront une excellente base d'orientation pour l'élaboration du programme de recherche.

À notre avis, il y a une importante conclusion à ne pas oublier compte tenu de l'appui impressionnant qu'a reçu le processus de consultation et des réactions positives qu'il a suscitées. Statistique Canada doit poursuivre l'effort d'information et d'initiation entrepris par cette consultation.

Renvoi

[1] Les indices provinciaux établis tous les trimestres sont distribués annuellement pour toutes les provinces (excluant les centres urbains de moins de 30,000 habitants). Toutefois, leur diffusion est limitée aux points de contact de Statistique Canada, à des fins d'utilisation par les bureaux provinciaux de la statistique.

AN OVERVIEW OF THE PAPERS ON PRICE LEVEL MEASUREMENT

Pour vous fournir une version dans la langue officielle de votre choix, le texte anglais est suivi du texte français (p.71) dans cette publication.

W.E. Diewert

The Statistics Canada Conference on Price Measurement was held in Ottawa on November 22-24, 1982. The paper by Baumgarten and Hodgins indicates the purpose of the conference and provides an introduction to most of the research issues that were discussed at the conference. Hence we will not review the background material presented in their paper but we will try to survey briefly the contents of the papers presented at the conference.

The conference papers have been organized into four groups: (i) the theoretical foundations of the CPI (Consumer Price Index), (ii) specific price measurement issues (which were raised in the Baumgarten-Hodgins paper), (iii) the measurement of price level change from the perspective of producer theory (rather than consumer theory) and (iv) business cycle issues.

The first theoretical paper is by Pollak, ‘‘The Theory of the Cost-of-Living Index’’. This paper was originally issued by the U.S. Bureau of Labor Statistics as BLS Working Paper 11 in June 1971. Due to its length, it was never formally published until the present. However, the paper has been extremely influential and is regarded as a ‘‘classic’’ in the field by researchers in index number theory. Pollak not only develops the theory of the cost-of-living index which the CPI approximates, he also studies the related problem of measuring welfare. Some of the contributions that Pollak makes in this paper are: (i) he provides a systematic theory of empirically implementable bounds for the single consumer cost-of-living index, (ii) he systematically develops the theory of the Malmquist [1953] quantity or welfare index, and (iii) he brings out the connection between functional forms for a consumer’s welfare or utility function and the functional form for the corresponding cost-of-living index.

The paper by Diewert, ''The Theory of the Cost-of-Living Index and the Measurement of Welfare Change'', continues to develop the theoretical foundations of CPI. Drawing on Pollak [1981], Diewert develops the many- consumer theory of the cost-of-living index. Diewert adapts a technique due originally to Konüs [1924] and shows that many theoretically unobservable indexes may be bounded by empirically computable Paasche and Laspeyres indexes. He also discusses several other issues from a theoretical perspective, including: (i) the use of the chain principle versus the fixed base principle in constructing index numbers, (ii) the problems involved in constructing group welfare indexes, (iii) the construction of subindexes and how they may be combined to form an overall index and (iv) the treatment of housing and other consumer durables in the CPI.

The paper by Jorgenson and Slesnick, **''Individual and Social cost of Living Indexes''**, is an intellectual tour de force as Russell notes in his perceptive comments on the paper. It is not possible to summarize accurately the entire paper in this brief introduction; however, some of the major ideas in the paper can be mentioned here. The core of the Jorgenson-Slesnick paper is an econometric model of aggregate consumer behavior which is based on Lau's [1977] theory of exact aggregation which in turn (greatly) generalizes the aggregate model of consumer behavior implemented by Berndt, Darrough and Diewert [1977]. The basic idea is that each household is assumed to have the same preferences (apart from a stochastic term) once we standardize for various demographic characteristics or attributes such as family size, age of head, region of residence, race, rural versus urban, etc. By shrewdly choosing the functional form for the household preferences, Jorgenson and Slesnick end up with a system of aggregate consumer demand functions that depend on consumer prices, the distribution of income, and the distribution of household characteristics. Given aggregate time series and some cross section data the parameters which characterize the household preferences may be estimated econometrically and this is done in Jorgenson, Lau and Stoker [1982]. Once household preferences are known, individual cost-of-living or welfare indexes may be calculated as well as aggregate indexes. Jorgenson and Slesnick also compare their ''econometric'' approach to the construction of aggregate cost-of-living and social welfare indexes to the ''exact'' index number approach outlined in the earlier paper by Diewert. There are advantages and disadvantages to each approach. One disadvantage of the Jorgenson-Slesnick approach is that new time series information would lead to new econometric estimates of the household preference parameters and hence new

historical index numbers would have to be calculated. However, there is much to be praised in the Jorgenson-Slesnick approach as Russell notes in his comment.

Riddell in his paper, "Leisure Time and the Measurement of Economic Welfare" extends the traditional theory of the cost-of-living index to incorporate the consumer-worker's labour supply (or leisure choice) decision. He also empirically implements his approach using Canadian per capita data. He compares his approach with the traditional approach (which neglects the labour supply decision) and finds that real income growth is dramatically slower in Canada when the labour supply decision is taken into account. It should also be mentioned that Riddell's approach takes into account income taxes as well as commodity taxes. Thus if the government decreased commodity taxes and increased income taxes in a neutral manner, Riddell's welfare indicator would show no change whereas a "traditional" welfare indicator (such as per capita national income deflated by the CPI) would show an increase. Riddell also extends and empirically implements the Pencavel [1977] – Cleeton [1982] theory of real wage indexes.

In the paper, "Preference Diversity and Aggregate Economic Cost-of-Living Indexes", Blackorby and Donaldson look for conditions on the preferences of individual consumers that are sufficient to ensure that an aggregate cost-of-living index is independent of the distribution of utilities or real incomes of the households in the economy. Their results are interesting but essentially negative: it appears that individual household preferences must be homothetic in order for the aggregate cost-of-living index to be independent of the distribution of utilities. Of course, homothetic preferences must be rejected on empirical grounds since homothetic preferences imply unitary income elasticities for all goods (which contradicts Engel's Law).

Eichhorn and Voeller in their paper, "The Axiomatic Foundations of Price Indexes and Purchasing Power Parities", outline the "axiomatic" or "test" approach to index number theory, as opposed to the "economic" approach which was followed by the earlier papers from Pollak to Blackorby and Donaldson. In his comments on Eichhorn and Voeller, Blackorby explains the difference between the "test" approach and the "economic" approach to a price index: in the former approach, prices and quantities in both periods are taken to be independent variables whereas in the latter approach, quantities are regarded as dependent variables. It should be mentioned that Diewert in his paper "The Theory

of Cost-of-Living Index and the Measurement of Welfare Change'' provides an axiomatic characterization of the ''economic'' approach to the cost-of-living index. In defence of Eichhorn and Voeller's approach, it should be noted that the ''economic'' approach to the cost-of-living index followed by most authors in this volume may not be the correct one. The ''economic'' approach to index number theory rests on the assumption of utility maximizing behaviour which may or may not be true. If it is not true, then there is a need for an alternative theoretical foundation for the CPI, and the ''axiomatic'' approach developed by Eichhorn and Voeller can fill this need.

The second group of papers deals with the specific price level measurement issues raised in the Baumgarten-Hodgins paper and we now turn to a brief description of these papers.

The paper by Triplett entitled, ''Escalation Measures: What is the Answer? What is the Question?'' asks the following question: what change in some economic variables is necessary to compensate for inflation? Examples of economic variables are: (i) expenditures, (ii) income, (iii) wealth, and (iv) a price such as a wage rate (see the papers by Riddell and Gillingham and Greenlees for examples of different types of cost-of-living indexes). Triplett's paper points out the pitfalls inherent in using the CPI as an escalator in inappropriate situations. Schaefer's comment on the paper raises some mild objections to the tenor of Triplett's arguments. Schaefer and Tripplet also discuss the problem of distinguishing between imported and domestic inflation, or put another way, how can the contribution of changes in the terms of trade to the domestic inflation rate be measured? An answer to this question is provided by Diewert in the paper, ''The Theory of Output Price Index and the Measurement of Real Ouput Change''.

In his paper, ''Impact of the Choice of Formulae on the Canadian Consumer Price Index'', Généreux compares the Paasche, Laspeyres and Fisher indexes using Canadian data. Most official CPIs are of the fixed base Laspeyres type and hence if they are regarded as approximations to a cost-of-living index, they will be biased upwards because they neglect substitution effects (the substitution of relatively lower priced goods for more expensive goods). This substitution bias may be minimized by using the chain principle (i.e., by changing the base basket of goods more frequently) or by using a ''superlative'' index number formula such as the Fisher [1922] ideal formula which will take the substitution phenomenon into account. Généreux finds surprisingly little substitution bias in the Canadian data.

In “Linking Price Index Numbers”, Szulc develops some theoretical formulae which enables one to analyze the difference between fixed base and chain indexes.

In her paper, ‘‘The Treatment of Housing in a Cost-of-Living Index: Rental Equivalence and User Cost’’, Darrough discusses one of the most vexing problems in the construction of a CPI: namely, how to calculate the price of housing. She compares and contrasts the two leading approaches to the problem (from the viewpoint of economic theory): the rental equivalence and the user cost approach. The two approaches and their advantages and disadvantages are clearly explained in the paper. One of the most important points to emerge from her paper (and the earlier paper by Diewert in this volume) is that the tax treatment of housing plays an important role in the construction of the user cost of owner-occupied housing and the rental equivalence approach to modelling the price of housing may not be able to capture this role.

Another paper which emphasizes tax considerations is ‘‘The Incorporation of Direct Taxes into a Consumer Price Index’’ by Gillingham and Greenlees. The authors construct a Tax and Price Index (TPI) that conceptually measures the before-tax income which is necessary to yield an after-tax income equal to the expenditure required to purchase a fixed basket of consumption goods. The reader should compare the Gillingham-Greenlees TPI with the minimum non-labour income indexes suggested by Riddell. Riddell’s indexes also take taxes into account, but in addition, substitution phenomenon are taken into account. A problem with the Gillingham-Greenlees framework is that a household with low wages and high hours could be treated in the same manner as a household with high wages and low hours worked. Nevertheless, the Gillingham-Greenlees paper is an impressive piece of theoretical and empirical work.

In his paper, ‘‘Public Goods and Price Indexes’’, Montmarquette provides an interesting theoretical and empirical treatment of another vexing problem in the construction of a CPI: namely, how to take into account the consumption of ‘‘public goods’’ provided by the government, such as parks, roads, public television broadcasting, etc. Montmarquette implements his theoretical model (based on the Borcherding and Deacon [1972] median voter model) using Canadian data.

An example of a good which is not a pure public good nor a pure private good (in Canada) is medical services. In their paper, ''Prices, Proxies and Productivity: An Historical Analysis of Hospital and Medical Care in Canada'', Barer and Evans provide a comprehensive survey of the data on medical services in Canada. They also discuss various theoretical approaches to the modelling of the demand for medical services from the perspective of economic theory. Not all of the analysis presented by Barer and Evans is free from controversy as the comment by Greenlees and Manser indicates.

In ''Regional Price Indexes: The Canadian Practice and Some Potential Extensions'', Denny and Fuss discuss the theoretical and practical problems involved in comparing price levels between regions from the perspective of the ''economic'' approach to index number theory. Eichhorn and Voeller on the other hand address the regional comparisons problem from the perspective of the ''test'' or ''axiomatic'' approach to index number theory. Denny and Fuss also provide some empirical evidence about regional price differences in Canada.

In ''Quality Adjustment, Hedonics and Modern Empirical Demand Analysis'', Berndt attempts to provide empirically implementable techniques for dealing with the quality change problem. His analysis skillfully blends the hedonic techniques of Griliches [1971] and Lau [1982] along with modern capital theory. In his comment on the Berndt paper, Triplett takes issue with some of Berndt's analysis and presents his own views on how quality change should be modelled.

As is noted in the Statistics Canada [1982; p.105] reference paper on the consumer price index (which is perhaps the best background paper for the material in this volume), the treatment of seasonal commodities is probably one of the most controversial issues in the construction of CPI or cost-of-living indexes. A seasonal commodity is one that is available only at certain times of the year (e.g., certain kinds of fruits and vegetables, downhill skiing services, professional sports services, etc.) or one where either the price or quantity sold varies considerably at different times of the year (e.g., bathing suits, the consumption of alcoholic beverages, Christmas gifts, etc.). Thus seasonal commodities are goods or services whose supplies vary with the time of year (with changing weather conditions being the principle cause of these supply changes) or whose demands vary with the time of year (due to custom or weather influences). From the perspective of constructing a monthly

CPI, the existence of seasonal commodities creates at least two difficulties: (i) at certain seasons, a commodity may simply be unavailable and hence no price quotation for it may exist, and (ii) even if price and quantity data are available for each season, it is difficult to obtain a ''basket'' of seasonal goods that is representative for all seasons; hence the final basket Laspeyres approach to the construction of ''monthly'' price indexes does not give meaningful month-to-month comparisons of prices, since normally when a seasonal price increases, the quantity purchased will decline (i.e., the ''basket'' will change). There are three papers in this volume that deal with aspects of the seasonal problem and we shall discuss each briefly.

The paper ''Measuring the Current Rate of Inflation'' by D. Rhoades and N. Elhawary-Rivet may be viewed as an introduction to the complexities of constructing a monthly CPI that is somehow consistent with an annual CPI. The authors compare a **monthly** CPI (which compares the level of prices in a given month with the preceding month) with an **annual** CPI (which compares the level of prices in a given month with the same month in the preceding year). They argue that the annual inflation rate is a smoothed version of the monthly rates and that the annual rate lags behind corresponding monthly rates by five to six months. Thus given that policymakers wish to know what the annual inflation rate for 1983 will be, the monthly inflation rate for say July (multiplied by 12) may give us an approximation to the annual 1983 rate well in advance of the end of the year. However, the problem with the monthly rates is that they are more erratic; i.e., the annual rate (since it may be viewed as a product of the monthly rates) tends to be ''smoother''.Rhoades and Elhawary-Rivet go on to discuss various methods (including one proposed by Moore which is discussed by him in his paper later in this volume) for forecasting an annual inflation rate using current monthly inflation rates. The discussant to their paper, Browne, summarizes the paper in a particularly clear fashion and then goes on to raise several important issues: (i) is the CPI the best measure of annual price change in a country (i.e., what about the Implicit Price Deflator for Gross National Expenditure or the Industry Selling Price Indexes as measures of inflation), (ii) should seasonally adjusted data be used in order to construct an annual inflation rate, and (iii) how can we construct ''core'' or ''nontransitory'' measures of inflation? None of these questions can be answered in a simple manner. Question (i) is discussed by Diewert in ''The Theory of the Output Price Index and the Measurement of Real Output Change''. The answer to question (ii) is more complex and hinges on what definition of the annual inflation rate that we wish to adopt and what

precisely we want a seasonal adjustment procedure to accomplish. Finally, the theory of subindexes due to Pollak [1975] and discussed in the earlier paper by Diewert is relevant to answering question (iii).

In their paper "The Estimation of Seasonal Variations in Consumer Price Indexes", Dagum and Morry ask the seasonal adjustment procedure to produce a "smooth" series and they also ask that revisions to currently seasonally adjusted series be small. The authors go on to summarize various methods that have been suggested in order to seasonally adjust series such as regression methods, moving average methods, and ARIMA (autoregressive integrated moving average) methods. They also discuss whether the micro data should be directly seasonally adjusted or whether the data should be seasonally adjusted at the aggregate level.

The starting point in Diewert's paper, "The Treatment of Seasonality in a Cost-of-Living Index", is an observation due perhaps to Richard Stone [1956, p.75]: there is no seasonality problem if we always compare the current year's prices and quantities with the corresponding prices and quantities of a base year. Thus annual price indexes can be formed using traditional index number formulae provided that we regard each commodity in each season as a separate commodity. Diewert then observes that there is no reason why we should restrict our comparisons to the base year's January to December with the current year's January to December observations: any consecutive 12 months can be compared with the price and quantity data in the corresponding 12 months in the base year. This leads to a monthly series of annual inflation rates, giving the level of prices for the year which ends at that month relative to the prices of a base year. Month-to-month changes in this series may be extrapolated to yield estimates for future annual inflation rates. Thus Diewert's monthly series of annual inflation rates is a series from which the seasonal elements have been removed and it may be used to forecast future annual inflation rates if this is desired. In effect, Diewert's proposed method of seasonal adjustment seems to provide a relatively simple solution to the problem of dealing with seasonal goods in the context of the CPI.

Section IV in this volume deals with the theoretical foundations of: (i) real output indexes, (ii) output price indexes of GNP price deflators and (iii) total factor productivity indexes. Section IV contains only one paper, "The Theory of the Output Price Index and the Measurement of Real Output Change", by Diewert. Many years ago, Hicks [1940]

and Samuelson [1950] discussed and contrasted the measurement of growth and inflation from either the producer perspective or the consumer perspective. Hicks [1975, p.317] noted under what conditions the two methods of measurement can lead to the same answers. The conditions for equivalence are not satisfied in any realistic economy and hence it is useful to consider both perspectives. Thus Diewert discusses growth and inflation measurement from the producer perspective in the present paper. In particular, the Fisher and Shell [1972], Samuelson and Swamy [1974] and Archibald [1977] concepts of output price indexes are discussed as is the Hicks [1981, p.256] and Caves, Christensen and Diewert [1982] concept of a total factor productivity index. Diewert also uses the producer framework in order to derive some terms of trade adjustment indexes which estimate what proportion of an open economy's private production sector growth between two periods is due to changes in the terms of trade (changes in the prices of exports relative to the prices of imports).

Section V in this volume is entitled, ''Price Behaviour and Business Cycles''. The first paper in this section by Moore, ''The Consumer Price Index and Signals of Recession and Recovery'', deals with some interesting measurement problems: (i) how can we determine when a recession starts (and ends), and (ii) how can we forecast the start (or end) of a recession? Moore applies his measurement techniques to Canada and contrasts the Canadian and U.S. recent experience.

The second paper in the Business Cycle section is ''Price Behaviour and Economic Prospects'' by Jay Forrester. This paper does not deal with the rather narrow measurement issues that are the concern of the other papers in this volume, but rather it deals with the fundamental economic and institutional causes of growth (or the lack of it) in output and prices. Forrester draws on his research over the years to present us with a comprehensive theory of the business cycle. He concludes his paper with some recommendations for future research.

Martin Wilk, the Chief Statistician for Canada, offers some closing comments to conclude the volume.

References to papers not in this volume

Archibald, R.B. [1977], "On the Theory of Industrial Price Measurement: Output Price Indexes", *Annals of Economic and Social Measurement* 6, pp.57-72.

Berndt, E.R., W.E. Diewert and M.N. Darrough [1977], "Flexible Functional Forms and Expenditure Distributions: An Application to Canadian Consumer Demand Functions", *International Economic Review* 18, pp.651-675.

Borcherding, T.E. and R.T. Deacon [1972], "The Demand for Services of Non Federal Governments", *American Economic Review* 62, pp.891-901.

Caves, D.W., L.R. Christensen and W.E. Diewert, [1982], "The Economic Theory of Index Numbers and the Measurement of Input, Output and Productivity", *Econometrica* 50, pp.1393-1414.

Cleeton, D.L. [1982], "The Theory of Real Wage Indices", *American Economic Review* 72, pp.214-225.

Fisher, F.M. and K. Shell [1972], "The Pure Theory of the National Output Deflator", pp.49-113 in *The Economic Theory of Price Indices* New York: Academic Press.

Fisher, I. [1922], *The Making of Index Numbers*, Boston: Houghton Mifflin Company.

Griliches, Zvi [1971], "Introduction: Hedonic Prices Revisited", pp.3-15 in *Price Indexes and Quality Change*, Zvi Griliches (ed.), Cambridge: Harvard University Press.

Hicks, J.R. [1940], "The Valuation of the Social Income", *Economica* 7, pp.105-124.

__________ [1975], "The Scope and Status of Welfare Economics", *Oxford Economic Papers* 27, pp.307-326.

__________ [1981], *Wealth and Welfare*, Cambridge: Harvard University Press.

Lau, L.J. [1977], ''Existence Conditions for Aggregate Demand Functions'', Technical Report No. 248, Institute for Mathematical Studies in the Social Sciences, Stanford University, Stanford, California.

_________ [1982], ''The Measurement of Raw Material Inputs'', pp.167-200 in V. Kerry Smith and J.V. Krutilla (eds.), *Explorations in Natural Resource Economics*, Baltimore: Johns Hopkins Press.

Jorgenson, D.W., L.J. Lau and T.M. Stoker [1982], ''The Transcendental Logarithmic Model of Aggregate Consumer Behavior'', pp.97-238 in *Advances in Econometrics*, Vol. 1, R.L. Basmann and G.F. Rhodes Jr. (eds.), Greenwich: JAI Press.

Konüs, A.A. [1924], ''The Problem of the True Index of the Cost-of-Living'', Translated in *Econometrica* 7 [1939], pp.10-29.

Malmquist, S. [1953], ''Index Numbers and Indifference Surfaces'', *Trabajos de Estadistica* 4, pp.209-242.

Pencavel, J.H. [1977], ''Constant Utility Index Numbers of Real Wages'', *American Economic Review* 67, pp.91-100.

Pollak, R.A. [1975], ''Subindexes of the Cost-of-Living Index'', *International Economic Review* 16, pp.135-150.

_________ [1981], ''The Social Cost-of-Living Index'', *Journal of Public Economics* 15, pp.311-336.

Samuelson, P.A. [1950], ''The Evaluation of Real National Income'', *Oxford Economic Papers* 2, pp.1-29.

_________ and S. Swamy [1974], ''Invariant Economic Index Numbers and Canonical Duality: Survey and Synthesis'', *American Economic Review* 64, pp.566-593.

Statistics Canada [1982], "The Consumer Price Index Reference Paper: Concepts and Procedures", Catalogue No. 62-553, Occasional, Ottawa, Canada.

Stone, R. [1956], *Quarterly and Price Indexes in National Accounts* Paris: The Organization for European Economics Co-operation.

APERÇU DES COMMUNICATIONS SUR LA MESURE DE LA VARIATION DES PRIX

To provide you with a version in the official language of your choice, the French text is preceeded by the English text (p.59) in this publication.

W.E. Diewert

La Conférence de Statistique Canada sur la mesure de la variation des prix a eu lieu à Ottawa du 22 au 24 novembre 1982. La communication de Baumgarten et Hodgins indique l'objet de la conférence et présente la plupart des problèmes de recherche qui y ont été débattus. Par conséquent, nous ne reprendrons pas les données de base présentées dans cette communication, mais nous ferons une brève revue du contenu des communications présentées à la conférence.

Nous avons divisé les communications en quatre groupes: (i) les fondements théoriques de l'IPC (indice des prix à la consommation), (ii) les problèmes particuliers de mesure de la variation des prix (soulevés dans la communication de Baumgarten-Hodgins), (iii) la mesure de la variation des niveaux de prix dans la perspective de la théorie des producteurs (plutôt que de la théorie des consommateurs) et (iv) les problèmes du cycle d'affaires.

La première communication théorique est celle de Pollak, ''La théorie de l'indice du coût de la vie''. Cette communication a été d'abord publiée par le U.S. Bureau of Labor Statistics, dans le *BLS Working Paper 11*, en juin 1971. En raison de sa longueur, ce texte n'avait jamais été officiellement publié avant maintenant. Cependant, il a eu une influence considérable et les chercheurs le considèrent aujourd'hui comme un ''classique'' dans le domaine de la théorie des nombres-indices. Pollak ne développe pas seulement la théorie de l'indice du coût de la vie, qui correspond à peu près à l'IPC, mais il étudie aussi le problème connexe de la mesure du bien-être. Dans cette communication, Pollak: (i) présente une théorie systématique des limites empiriquement applicables d'un indice du coût de la vie d'un consommateur unique, (ii) développe la théorie de l'indice Malmquist [1953] de

quantité ou du bien-être, et (iii) fait ressortir le lien entre les formes fonctionnelles du bien-être du consommateur ou sa fonction d'utilité et la forme fonctionnelle de l'indice correspondant du coût de la vie.

La communication de Diewert sur la théorie de l'indice du coût de la vie et de la mesure de la variation du bien-être poursuit le développement des fondements théoriques de l'IPC. S'inspirant de Pollak [1981], Diewert élabore la théorie de l'indice du coût de la vie de consommateurs multiples. Il adapte une technique dont l'ancêtre est Konüs [1924] et démontre que de nombreux indices théoriquement inobservables peuvent être limités par des indices empiriquement calculables de Paasche et de Laspeyres. Il examine également plusieurs autres questions dans une perspective théorique, notamment: (i) l'application du principe de l'enchaînement par opposition au principe de la base fixe dans la construction des nombres-indices, (ii) les problèmes que comportent la construction d'indices de bien-être collectif, (iii) la construction de sous-indices et la manière de les regrouper pour former un indice global, et (iv) le traitement du logement et des autres biens durables de consommation dans l'IPC.

La communication de Jorgenson et Slesnick, ''Indices individuel et social du coût de la vie'', est un tour de force intellectuel, comme Russell le note dans son commentaire perceptif sur cette communication. Il est impossible de résumer fidèlement l'ensemble de la communication dans cette brève introduction; cependant, on peut reprendre ici certaines des grandes orientations de leur texte. Le corps de la communication de Jorgenson et Slesnick est un modèle économétrique du comportement agrégatif des consommateurs fondé sur la théorie de l'agrégation exacte de Lau [1977] qui généralise (dans une grande mesure) le modèle agrégatif du comportement des consommateurs mis en oeuvre par Berndt, Darrough et Diewert [1977]. L'idée fondamentale est de supposer que chaque ménage a les mêmes préférences (sauf pour un terme stochastique) une fois que nous avons normaliser les ménages pour les divers attributs ou caractéristiques démographiques comme la taille de la famille, l'âge du chef, la région de résidence, la race, la catégorie d'habitat, etc. En choisissant avec courage une forme fonctionnelle pour les préférences des ménages, Jorgenson et Slesnick aboutissent à un système agrégé de fonctions de demande qui dépendent des prix à la consommation, de la distribution du revenu, et de la répartition des caractéristiques des ménages. À partir des séries chronologiques agrégatives et de certaines données

longitudinales, on peut estimer économétriquement les paramètres qui caractérisent les préférences des ménages et c'est ce que font Jorgenson, Lau et Stoker [1982]. Une fois connues les préférences des ménages, on peut calculer les indices individuels du coût de la vie ou du bien-être ainsi que des indices agrégatifs. Jorgenson et Slesnick comparent également leur approche ''économétrique'' de la construction d'indices agrégatifs du coût de la vie et du bien-être social à l'approche du nombre-indice ''exact'' déjà exposée dans le document de Diewert. Chaque approche comporte ses avantages et ses inconvénients. Un des inconvénients de l'approche de Jorgenson et Slesnick, c'est que les données de nouvelles séries chronologiques mèneraient à de nouvelles estimations économétriques des paramètres relatifs aux préférences des ménages et que, partant, il faudrait calculer de nouveaux nombres-indices chronologiques. Cependant, l'approche de Jorgenson et Slesnick ne manque pas d'attrait, comme le note Russell dans son commentaire.

Dans sa communication sur le temps de loisirs et la mesure du bien-être économique, Riddell étend la théorie traditionnelle de l'indice du coût de la vie de manière à y faire entrer la décision d'offre de main-d'oeuvre (ou le choix de loisirs) du consommateur-travailleur. Il applique aussi empiriquement son approche en utilisant des données per capita pour le Canada. Il compare son approche avec l'approche classique (qui ne tient pas compte de la décision d'offre de main-d'oeuvre) et constate que la croissance du revenu réel est sensiblement plus faible au Canada lorsque l'on tient compte de la décision d'offre de main-d'oeuvre. Il y a lieu de mentionner également que l'approche de Riddell prend en compte les impôts sur le revenu ainsi que les taxes sur les biens. Ainsi, si le gouvernement diminuait les taxes sur les biens et augmentait les impôts sur le revenu d'une manière neutre, l'indicateur de bien-être de Riddell ne ferait ressortir aucun changement, alors qu'un indicateur de bien-être ''classique'' (comme le revenu national par habitant dégonflé selon l'IPC) indiquerait une augmentation. Riddell étend aussi et applique empiriquement la théorie de Pencavel [1977] et de Cleeton [1982] des indices des salaires réels.

Dans leur communication, ''Diversité des préférences et indices économiques agrégatifs du coût de la vie'', Blackorby et Donaldson recherchent, à l'égard des préférences des consommateurs individuels, des conditions qui sont suffisantes pour assurer qu'un indice agrégatif du coût de la vie est indépendant de la distribution des utilités ou des revenus réels des ménages dans l'économie. Leurs résultats sont intéressants, mais essentiellement négatifs: il semble que les préférences des ménages individuels doivent être homothétiques

pour que l'indice agrégatif du coût de la vie soit indépendant de la distribution des utilités. Naturellement, il faut rejeter les préférences homothétiques pour des motifs empiriques puisqu'elles supposent des élasticités unitaires du revenu pour l'ensemble des biens (ce qui contredit la loi d'Engel).

Dans ''Bases axiomatiques des indices de prix et des parités de pouvoir d'achat'', Eichhorn et Voeller exposent l'approche ''axiomatique'' ou ''d'essai'' pour la théorie des nombres-indices, par opposition à l'approche ''économique'' adoptée dans les communications antérieures de Pollak, Blackorby et Donaldson. Dans son commentaire sur Eichhorn et Voeller, Blackorby explique la différence entre l'approche ''d'essai'' et l'approche ''économique'' dans le cas d'un indice de prix: dans la première, les prix et les quantités des deux périodes sont des variables indépendantes, alors que dans la seconde les quantités sont considérées comme des variables dépendantes. Il faut mentionner que Diewert, dans son texte sur la théorie de l'indice du coût de la vie et de la mesure de la variation du bien-être, caractérise axiomatiquement l'approche ''économique'' de l'indice du coût de la vie. Il faut dire, à la défense de l'approche d'Eichhorn et Voeller, que l'approche ''économique'' de l'indice du coût de la vie adoptée par la plupart des autres auteurs de ce volume n'est peut-être pas la bonne. L'approche ''économique'' de la théorie des nombres-indices part du principe d'un comportement qui maximise l'utilité, ce qui n'est pas forcément le cas. Si ce n'est pas le cas, il faut alors un autre fondement théorique pour l'IPC, et l'approche ''axiomatique'' élaborée par Eichhorn et Voeller peut combler ce besoin.

Le second groupe de communications traite de problèmes particuliers de la mesure du niveau des prix soulevés dans la communication de Baumgarten et Hodgins et nous passons maintenant à une brève description de ces communications.

La communication de Triplett intitulée ''Mesures d'indexation: Quelle est la réponse? Quelle est la question?'' pose la question suivante: quel changement faut-il apporter à une variable économique quelconque pour compenser l'inflation? Il donne comme exemples de variables économiques: (i) les dépenses, (ii) les revenus, (iii) la richesse, et (iv) un prix comme un taux salarial (voir les communications de Riddell et Gillingham et de Greenlees pour des exemples des différents types d'indices du coût de la vie). Triplett signale les dangers inhérents que comporte l'utilisation de l'IPC comme facteur d'indexation dans des cas inappropriés. Le commentaire de Schaefer sur sa communication soulève quelques légères

objections aux arguments qu'invoque Triplett. Schaefer et Triplett discutent également du problème de la distinction entre l'inflation importée et l'inflation interne, ou, autrement dit, sur la façon de mesurer l'apport des changements dans les termes d'échanges par rapport au taux d'inflation interne? Diewert répond à cette question dans sa communication sur ''la théorie de l'indice des prix des extrants et de la mesure de la variation de la production réelle''.

En traitant de ''l'impact du choix des formules sur l'Indice des prix à la consommation du Canada'', Généreux compare les indices de Paasche, Laspeyres et Fisher en utilisant des données canadiennes. La plupart des IPC sont du type Laspeyres à base fixe et, par conséquent, s'ils sont considérés comme des approximations d'un indice du coût de la vie, ils seront entachés d'un biais systématique vers le haut étant donné qu'ils ne tiennent pas compte des effets de substitution (la substitution de biens de prix relatifs plus faibles à des biens plus coûteux). Ce biais de substitution peut être minimisé par l'application du principe de l'enchaînement (c.-à-d. par une modification plus fréquente du panier de base) ou par l'utilisation d'une formule ''superlative'' de nombres-indices comme la formule idéale de Fisher [1922], qui tient compte du phénomène de substitution. Généreux décèle un biais de substitution qui étonne par sa faible importance dans les données canadiennes.

Dans "Enchaînement des indices de prix", Szulc développe certaines formules theoriques qui permettent d'analyser la différence entre les indices à base fixe et les indices en chaîne.

Dans son texte intitulé ''Le traitement du logement dans un indice du coût de la vie: équivalence locative et coût d'utilisation'', Darrough examine l'un des problèmes les plus difficiles de la construction d'un IPC, à savoir: comment calculer le prix du logement. Elle compare et met en contraste les deux principales approches du problème (du point de vue de la théorie économique): l'équivalence locative et le coût d'utilisation. Elle explique clairement les deux approches et les avantages et les inconvénients de chacune. L'un des points les plus importants qui ressort de son texte (et de la communication antérieure de Diewert dans le présent volume), c'est que le traitement fiscal du logement joue un rôle important dans la construction du coût d'utilisation des logements occupés par leur propriétaire, rôle que l'approche de l'équivalence locative pour la modélisation du prix du logement n'arrive peut-être pas à saisir.

Une autre communication qui met en lumière des considérations d'ordre fiscal est "Incorporation des impôts directs à un indice des prix à la consommation", par Gillingham et Greenlees. Les auteurs construisent un indice des impôts et des prix (IIP) qui mesure conceptuellement le revenu avant impôt qui est nécessaire pour obtenir un revenu après impôt égal à la dépense que suppose l'achat d'un panier fixe de biens de consommation. Le lecteur devrait comparer l'IIP de Gillingham-Greenlees avec les indices de revenu minimum ne provenant pas du travail que propose Riddell. Les indices de Riddell prennent en compte les impôts également, mais, en outre, ils tiennent compte du phénomène de substitution. Un des problèmes que pose le cadre de Gillingham-Greenlees, c'est qu'un ménage dont le revenu salarial est faible mais les heures de travail nombreuses peut être considéré sur le même pied qu'un ménage dont le revenu salarial est élevé et les heures de travail peu nombreuses. Néanmoins, la communication de Gillingham-Greenlees constitue un document théorique et empirique impressionnant.

Dans "Les biens publics et les indices de prix", Montmarquette présente un traitement théorique et empirique intéressant d'un autre problème complexe de la construction d'un IPC: comment tenir compte de la consommation des "biens publics" offerts par le gouvernement, comme les parcs, les routes, la télévision publique, etc. Montmarquette applique son modèle théorique (fondé sur le modèle de l'électeur médian de Borcherding et Deacon [1972]) avec des données canadiennes.

Les services médicaux constituent un exemple d'un bien qui n'est pas un bien public pur, ni un bien privé pur (au Canada). Dans leur communication sur "les prix, les mesures de substitution et la productivité (analyse historique des soins hospitaliers et médicaux au Canada)", Barer et Evans présentent une étude détaillée et complète des données sur les services médicaux au Canada. Ils discutent également de diverses approches théoriques de la modélisation de la demande de services médicaux dans la perspective de la théorie économique. L'analyse présentée par Barer et Evans n'est pas entièrement à l'abri de la controverse, comme l'indique le commentaire de Greenlees et Manser.

Dans leur texte sur "la pratique canadienne et certaines applications possibles des indices régionaux de prix", Denny et Fuss abordent les problèmes théoriques et pratiques que comporte la comparaison des niveaux de prix entre les régions dans la perspective de

l'approche ''économique'' de la théorie des nombres-indices. Eichhorn et Voeller, par ailleurs, abordent le problème des comparaisons régionales dans la perspective de l'approche ''d'essai'' ou ''axiomatique'' de la théorie des nombres-indices. Denny et Fuss offrent également certaines données empiriques sur les différences régionales de prix au Canada.

Dans sa communication sur ''l'ajustement de qualité, l'hédonique et l'analyse moderne empirique de la demande'', Berndt essaie de fournir des techniques empiriquement applicables pour aborder le problème du changement de qualité. Son analyse combine habilement les techniques hédoniques de Griliches [1971] et de Lau [1982] avec la théorie moderne du capital. Dans son commentaire sur la communication de Berndt, Triplett s'en prend à une partie de l'analyse de Berndt et présente ses propres vues sur la modélisation de la variation de qualité.

Comme l'indique le document de référence de Statistique Canada (1982; p.105) sur l'Indice des prix à la consommation (qui est peut-être le meilleur document de travail pour les textes réunis ici), le traitement des biens saisonniers est probablement l'une des questions les plus controversées dans la construction d'un IPC ou d'un indice du coût de la vie. Un bien saisonnier est un bien qui n'est disponible qu'à certaines époques de l'année (par ex., certains genres de fruits et de légumes, les services de ski alpin, les services de sport professionnel, etc.) ou un bien dont le prix ou la quantité varie considérablement selon l'époque de l'année (par ex., les maillots de bain, la consommation de boissons alcooliques, les cadeaux de Noël, etc.). Les biens saisonniers sont des biens ou des services dont l'approvisionnement varie avec l'époque de l'année (le changement des conditions météorologiques étant la cause principale de ces changements d'offre) ou dont la demande varie avec les saisons (sous l'influence des habitudes ou du temps). Dans la perspective de la construction d'un IPC mensuel, l'existence de biens saisonniers pose au moins deux difficultés: (i) à certaines saisons, un bien peut n'être tout simplement pas disponible, de sorte qu'il est impossible d'en relever le prix, et (ii) même s'il y a des données sur les prix et les quantités pour chaque saison, il est difficile d'obtenir un ''panier'' de biens saisonniers qui soit représentatif pour toutes les saisons; par conséquent, l'approche Laspeyres du panier final pour la construction des indices de prix ''mensuels'' ne se prête pas à des comparaisons mensuelles significatives des prix, puisque, normalement, lorsqu'un prix

saisonnier augmente, la quantité achetée diminue (c.-à-d., le ''panier'' change). Trois communications de ce volume traitent des aspects du problème saisonnier et nous les aborderons chacune brièvement.

La communication sur la mesure du taux courant d'inflation, par D. Rhoades et N. Elhawary-Rivet peut être considérée comme une introduction aux complexités de la construction d'un IPC mensuel qui soit à peu près convergent avec un IPC annuel. Les auteurs comparent un IPC **mensuel** (qui compare le niveau des prix d'un mois donné avec celui du mois précédent) avec un IPC **annuel** (qui compare le niveau des prix d'un mois donné avec celui du même mois de l'année précédente). Ils soutiennent que le taux annuel d'inflation est une version lissée des taux mensuels et que le taux annuel retarde de 5 ou 6 mois sur les taux mensuels correspondants. Ainsi, étant donné que les responsables des politiques veulent savoir quel sera le taux d'inflation annuel pour 1983, le taux d'inflation mensuel pour, mettons, juillet (multiplié par 12) peut donner une idée approximative du taux annuel de 1983 bien avant la fin de l'année. Cependant, le problème que posent les taux mensuels, c'est qu'ils sont plus erratiques, c.-à-d. que le taux annuel (puisqu'il peut être perçu comme un produit des taux mensuels) tend à être ''plus lisse''. Rhoades et Elhawary-Rivet poursuivent en discutant des diverses méthodes (y compris une méthode proposée par Moore, qu'il discute dans sa communication plus loin dans le présent volume) pour prévoir un taux annuel d'inflation à l'aide des taux mensuels courants. Dans son commentaire, Browne résume leur communication avec une clarté particulière, puis soulève plusieurs questions importantes: (i) l'IPC est-il la meilleure mesure de la variation annuelle des prix dans un pays (c.-à-d., que fait-on du déflateur implicite des prix pour la dépense nationale brute ou des indices des prix de vente dans l'industrie comme mesures de l'inflation), (ii) faut-il utiliser des données désaisonnalisées pour construire un taux d'inflation annuel et (iii) comment pouvons-nous construire des mesures ''centrales'' ou ''non transitoires'' de l'inflation? Il n'y a pas de réponse simple à ces questions. Diewert traite de la question (i) dans son texte sur la théorie de l'indice des prix de la production et de la mesure de la variation de la production réelle. La réponse à la question (ii) est plus complexe et est axée sur la définition du taux d'inflation annuel que nous voulons adopter et sur ce que nous voulons faire précisément avec une procédure de désaisonnalisation. Enfin, la théorie des sous-indices attribuable à Pollak [1975] et présentée dans la communication antérieure de Diewert est utile pour répondre à la question (iii).

Dans leur communication ‘‘L’estimation des variations saisonnières dans les indices de prix à la consommation’’, Dagum et Morry demandent à la procédure de désaisonnalisation de produire une série ‘‘lisse’’ et ils demandent également que les révisions des séries désaisonnalisées courantes soient faibles. Les auteurs résument ensuite diverses méthodes proposées pour désaisonnaliser des séries, par ex., les méthodes de régression, les méthodes de la moyenne mobile, et les méthodes de la moyenne mobile intégrée autorégressive. Ils se demandent également s’il y a lieu de désaisonnaliser directement les micro-données ou s’il faut les désaisonnaliser au niveau de l’ensemble.

Le point de départ de la communication de Diewert sur le ‘‘traitement de la saisonnalité dans un indice du coût de la vie’’ est une observation due peut-être à Richard Stone (1956, p.75): il n’y a pas de problème de saisonnalité si nous comparons toujours les prix et quantités de l’année courante avec les prix et quantités correspondants d’une année de base. Ainsi, on peut former des indices annuels de prix en utilisant les formules classiques de nombres-indices à la condition de considérer chaque bien, de chaque saison, comme bien distinct. Diewert fait ensuite observer qu’il n’y a pas de raison pour que nous restreignions nos comparaisons à la période janvier-décembre de l’année de base avec les observations de janvier-décembre de l’année en cours: on peut comparer toute période de 12 mois consécutifs avec les données sur les prix et les quantités des 12 mois correspondants de l’année de base. Cela donne une série mensuelle de taux annuels d’inflation, ce qui donne le niveau des prix pour l’année qui se termine ce mois-là par rapport aux prix d’une année de base. On peut extrapoler les variations d’un mois à l’autre de cette série pour obtenir des estimations des taux annuels futurs d’inflation. Ainsi, la série mensuelle de Diewert de taux annuels d’inflation est une série dont on a éliminé les éléments saisonniers et elle peut servir à prévoir les taux annuels futurs d’inflation si on le désire. Effectivement, la méthode de désaisonnalisation que propose Diewert semble offrir une solution relativement simple au problème des biens saisonniers dans le contexte de l’IPC.

La section IV du présent volume traite des fondements théoriques: (i) des indices de production réelle, (ii) des indices de prix de la production ou des déflateurs des prix du PNB, et (iii) des indices de la productivité des facteurs. La section IV ne contient qu’un seul document, qui traite de ‘‘la théorie de l’indice des prix de la production et de la mesure de la variation de la production réelle’’, par Diewert. Il y a de nombreuses années, Hicks [1940]

et Samuelson [1950] ont examiné et mis en contraste la mesure de la croissance et de l'inflation dans la perspective du producteur ou dans celle du consommateur. Hicks (1975, p.317) indique dans quelles conditions les deux méthodes de mesure peuvent conduire aux mêmes conclusions. Les conditions d'équivalence ne sont pas satisfaites dans une économie réaliste et, partant, il est utile de considérer les deux perspectives. Ainsi, Diewert discute dans la présente communication de la mesure de la croissance et de l'inflation dans la perspective du producteur. En particulier, les concepts de Fisher et Shell [1972], de Samuelson et Swamy [1974] et d'Archibald [1977] sur les indices de prix de la production sont discutés, tout comme le concept de l'indice de la productivité des facteurs de Hicks (1981, p.256) et de Caves, Christensen et Diewert [1982]. Diewert fait également appel au cadre du producteur afin de construire certains indices de rajustement des termes d'échange pour estimer quelle proportion de la croissance du secteur de la production privée d'une économie libre d'une période à l'autre est attribuable aux variations des termes d'échange (changements des prix des exportations par rapport aux prix des importations).

La section V du présent volume est intitulée, ''Comportement des prix et cycles économiques''. La première communication de cette section, ''L'indice des prix à la consommation et indicateurs de récession et de reprise'', par Moore, aborde certains problèmes intéressants de mesure: (i) comment pouvons-nous déterminer le début (et la fin) d'une récession, et (ii) comment pouvons-nous prévoir le début (ou la fin) d'une récession? Moore applique au Canada ses techniques de mesure et met en contraste l'expérience récente du Canada et des États-Unis.

La deuxième communication de la section des cycles économiques s'intitule ''Comportement des prix et perspectives économiques'', par Jay Forrester. Cette communication ne traite pas des problèmes plutôt étroits de mesure qui font l'objet des autres communications de ce volume, mais aborde plutôt les causes économiques et institutionnelles fondamentales de la croissance (ou de l'absence de croissance) de la production et des prix. Forrester puise aux recherches qu'il a effectuées au cours des ans pour nous présenter une théorie globale du cycle d'affaires. Il conclut avec quelques recommandations de recherche future.

Martin Wilk, Statisticien en chef du Canada, fait quelques commentaires de clôture pour terminer le volume.

Références

Archibald, R.B. [1977], "On the Theory of Industrial Price Measurement: Output Price Indexes", *Annals of Economic and Social Measurement* 6, pp.57-72.

Berndt, E.R., W.E. Diewert, et M.N. Darrough [1977], "Flexible Functional Forms and Expenditure Distributions: An Application to Canadian Consumer Demand Functions", *International Economic Review* 18, pp.651-675.

Borcherding, T.E. et R.T. Deacon [1972], "The Demand for Services of Non Federal Governments", *American Economic Review* 62, pp.891-901.

Caves, D.W., L.R. Christensen, et W.E. Diewert [1982], "The Economic Theory of Index Numbers and the Measurement of Input, Output and Productivity", *Econometrica* 50, pp.1393-1414.

Cleeton, D.L. [1982], "The Theory of Real Wage Indices", *American Economic Review* 72, pp.214-225.

Fisher, F.M. et K. Shell [1972], "The Pure Theory of the National Output Deflator", pp.49-113 dans *The Economic Theory of Price Indices* par F.M. Fisher et K. Shell, New York: Academic Press.

Griliches, Zvi [1971], "Introduction: Hedonic Prices Revisited", pp.3-15 dans *Price Indexes and Quality Change*, Zvi Griliches (ed.), Cambridge: Harvard University Press.

Hicks, J.R. [1940], "The Valuation of the Social Income", *Economica* 7, pp.105-124.

__________ [1975], "The Scope and Status of Welfare Economics", *Oxford Economic Papers* 27, pp.307-326.

[1981], *Wealth and Welfare*, Cambridge: Harvard University Press.

Jorgenson, D.W., L.J. Lau et T.M. Stoker [1982], ''The Transcendental Logarithmic Model of Aggregate Consumer Behavior'', pp.97-238 dans *Advances in Econometrics*, Vol. 1, R.L. Basmann et G.F. Rhodes Jr. (éd.), Greenwich: JAI Press.

Konüs, A.A. [1924], ''The Problem of the True Index of the Cost of Living'', traduit dans *Econometrica* 7 [1939], pp.10-29.

Lau, L.J. [1977], ''Existence Conditions for Aggregate Demand Functions'', Technical Report No. 248, Institute for Mathematical Studies in the Social Sciences, Stanford University, Stanford (Californie).

________ [1982], ''The Measurement of Raw Material Inputs'', pp.167-200 dans V. Kerry Smith et J.V. Krutilla (éd.) *Explorations in Natural Resource Economics*, Baltimore: Johns Hopkins Press.

Malmquist, S. [1953], ''Index Numbers and Indifference Surfaces'', *Trabajos de Estadistica* 4, pp.209-242.

Pencavel, J.H. [1977], ''Constant Utility Index Numbers of Real Wages'', *American Economic Review* 67, pp.91-100.

Pollak, R.A. [1975], ''Subindexes of the Cost of Living Index'', *International Economic Review* 16, pp.135-150.

________ [1981], ''The Social Cost of Living Index'', *Journal of Public Economics* 15, pp.311-336.

Samuelson, P.A. [1950], ''The Evaluation of Real National Income'', *Oxford Economic Papers* 2, pp.1-29.

________ et S. Swamy [1974], ''Invariant Economic Index Numbers and Canonical Duality: Survey and Synthesis'', *American Economic Review* 64, pp.566-593.

Statistique Canada [1982], ''Document de référence de l'Indice des prix à la consommation: concepts et procédés, n-b0- 62-553 au catalogue, hors série, Ottawa (Canada).

Stone, R. [1956], *Quarterly and Price Indexes in National Accounts*, Paris: Organization for European Economic Co-operation.

SECTION II

The Theoretical Foundations of a Consumer Price Index

Les fondements théoriques de l'Indice des prix à la consommation

THE THEORY OF THE COST-OF-LIVING INDEX

*Robert A. Pollak**
Department of Economics
University of Pennsylvania

SUMMARY

In this paper I summarize the theory of the "cost-of-living index" and the closely related theory of the "preference field quality index". The first section is devoted to notation and preliminary background, and the second summarizes the theory of the cost-of-living index. The next two sections discuss upper and lower bounds for the index, and the preference orderings for which these bounds are attained. In Section 5 I examine the effect on the index of the choice of a base indifference curve. Section 6 examines the form of the cost-of-living index corresponding to various specific preference orderings. In Sections 7 and 8 I examine the preference field quantity index, a quantity index which is conceptually similar to the cost-of-living index. I show that it is equal to the expenditure index deflated by the cost-of-living index if and only if the preference ordering is homothetic to the origin. In Section 9 I discuss the use of price indexes in empirical demand analysis and argue that the cost-of-living index is inappropriate for this purpose.

RÉSUMÉ

Dans ce document, je résume la théorie de "l'indice du coût de la vie" et la théorie étroitement liée de "l'indice de qualité du champ des préférences". Le premier chapitre expose la notation et des notions générales préliminaires, tandis que le deuxième résume la théorie

de l'indice du coût de la vie. Les deux chapitres suivants traitent des bornes supérieure et inférieure de l'indice et des classements des préférences pour lesquels ces bornes sont atteintes. Au chapitre 5, j'étudie l'effet du choix d'une courbe d'indifférence de base sur l'indice. Le chapitre 6 est consacré à l'étude de la forme de l'indice du coût de la vie correspondant à divers classements précis des préférences. Dans les chapitres 7 et 8, j'étudie l'indice de quantité du champ des préférences, un indice quantitatif, qui, sur le plan conceptuel, est analogue à l'indice du coût de la vie. Je montre qu'il est égal à l'indice des dépenses, "dégonflé" par l'indice du coût de la vie uniquement si le classement des préférences est homothétique par rapport à l'origine. Au chapitre 9, j'étudie l'utilisation des indices de prix dans l'analyse empirique de la demande et soutiens que l'indice du coût de la vie ne convient pas à cette fin.

Il existe un certain nombre de questions d'intérêt pratique et théorique que je n'aborde pas dans ce document. Il n'est pas fait mention des changements de qualité, des biens fournis par les pouvoirs publics ni du traitement de l'environnement. On suppose que les biens sont parfaitement divisibles et les coûts de transaction nuls. Les aspects intertemporels du problème d'allocation du consommateur ne sont pas pris en compte, de sorte que les sujets de l'épargne, des taux d'intérêt et des biens durables de consommation ne sont pas traités. J'ai laissé ces questions de coté non seulement parce qu'elles sont délicates, mais aussi parce qu'une critique théorique systématique dans l'un de ces domaines doit s'appuyer sur une connaissance approfondie de la théorie fondamentale. L'un des principaux objectifs de ce document est de présenter un tableau rigoureux de cette théorie. La matière des deux premiers chapitres devrait être connue, mais une bonne partie du contenu des chapitres qui suivent est nouvelle.

L'exposé consacré aux bornes de l'indice du coût de la vie, des chapitres 3 et 4, souligne l'importance des deux bornes, supérieure et inférieure. Il est bien connu que l'indice de Laspeyres constitue une borne supérieure de l'indice du coût de la vie basé sur une courbe d'indifférence obtenue dans la situation de référence, mais l'on n'a guère porté d'atttention à la borne inférieure correspondante. Au chapitre 3, j'établis les bornes supérieure

et inférieure, avant de caractériser, au chapitre 4, les classements des préférences pour lesquels ces bornes sont atteintes.

Il est bien connu que l'indice du coût de la vie n'est indépendant de la courbe d'indifférence de base que si le diagramme d'indifférence est homothétique par rapport à l'origine, et uniquement dans ce cas. Au chapitre 5, j'envisage plusieurs catégories de classements de préférences non homothétiques et étudie leurs conséquences pour le rapport entre l'indice du coût de la vie et la courbe d'indifférence de base.

Au chapitre 6, j'étudie une série de classements des préférences suffisamment simples pour convenir à des travaux empiriques et j'examine l'indice du coût de la vie correspondant à chacun d'entre eux. Étant donné que le calcul d'un indice du coût de la vie dépend de l'estimation d'un système d'équations de demande, il importe d'avoir une certaine idée des possibilités.

Au chapitre 7, je suis Malmquist (1953) dans la définition d'un indice de quantité du champ des préférences. Au chapitre 8, je montre que l'utilisation d'un indice du coût de la vie pour "dégonfler" l'indice des dépenses ne produit pas cet indice, sauf si le classement des préférences est homothétique par rapport à l'origine. Cela porte à croire que, si l'on veut un indice quantitatif, on doit essayer de l'obtenir directement plutôt qu'en divisant la dépense par un indice de prix.

L'exposé consacré, au chapitre 9, à l'utilisation des indices de prix dans l'analyse empirique de la demande laisse à penser que la pratique courante consistant à dégonfler les prix et les revenus par un indice de Laspeyres n'a aucun fondement théorique. Il implique également qu'une déflation de ces valeurs par l'indice du coût de la vie ne serait guère meilleure. De plus, comme on ne peut calculer l'indice du coût de la vie avant d'avoir estimé les paramètres inconnus du système d'équations de demande, une méthode qui exigerait une déflation des prix et des revenus par cet indice ne présenterait guère de valeur pratique.

In this paper I summarize the theory of the "cost-of-living index" and the closely related theory of the "preference field quality index". The first section is devoted to notation and preliminary background, and the second summarizes the theory of the cost-of-living index. The next two sections discuss upper and lower bounds for the index, and the preference orderings for which these bounds are attained. In Section 5 I examine the effect on the index of the choice of a base indifference curve. Section 6 examines the form of the cost-of-living index corresponding to various specific preference orderings. In Sections 7 and 8 I examine the preference field quantity index, a quantity index which is conceptually similar to the cost-of-living index. I show that it is equal to the expenditure index deflated by the cost-of-living index if and only if the preference ordering is homothetic to the origin. In Section 9 I discuss the use of price indexes in empirical demand analysis and argue that the cost-of-living index is inappropriate for this purpose.

There are a number of areas of practical and theoretical importance which I do not discuss in this paper. No mention is made of quality change, goods provided by governments, or the treatment of the environment. Goods are assumed to be perfectly divisible and transaction costs to be zero. Intertemporal aspects of the consumer's allocation problem are ignored, so the treatment of saving, interest rates, and consumer durables is not discussed. I have ignored these areas not only because they are difficult but also because a systematic theoretical attack on any of them must be based on a thorough understanding of the basic theory. A major purpose of this paper is to present a rigorous survey of that theory. The material in the first two sections should be familiar, but much of the material in later sections is new.

The discussion of bounds on the cost-of-living index in Sections 3 and 4 emphasizes the importance of both upper and lower bounds. It is well known that the Laspeyres index provides an upper bound to the cost-of-living index based on the indifference curve attained in the reference situation, but little attention has been given to the corresponding lower bounds. In Section 3 I establish upper and lower bounds, and in Section 4 I characterize the preference orderings for which these bounds are attained.

It is well known that the cost-of-living index is independent of the base indifference curve if and only if the indifference map is homothetic to the origin. In Section 5 I consider

several classes of non-homothetic preference orderings and examine their implications for the relationship between the cost-of-living index and the base indifference curve.

In Section 6 I consider a menu of preference orderings simple enough that they are suitable for empirical work and examine the cost-of-living index corresponding to each. Since computation of a cost-of-living index depends on estimation of an underlying system of demand equations, it is important to have some sense of what the possibilities are.

In Section 7 I follow Malmquist in defining a preference field quantity index. In Section 8 I show that use of a cost-of-living index to deflate the expenditure index does not yield this index unless the preference ordering is homothetic to the origin. This suggests that if one wants a quantity index, one must try to get it directly rather than by deflating expenditure by a price index.

The discussion in Section 9 of the use of price indexes in empirical demand analysis suggests that the common practice of deflating prices and income by a Laspeyres index has no theoretical standing. It also implies that deflating them by the cost-of-living index would be little better. Furthermore, since the cost-of-living index cannot be computed until the unknown parameters of the system of demand equations have been estimated, a procedure which required deflation of price and income by that index would be of little practical value.

The literature on the theory of the cost-of-living index is of uneven quality. The most lucid general treatments are those of Samuelson [1947], Wold and Jureen [1953], and Malmquist [1953]. A more recent survey is Afriat [1972]. Fisher and Shell [1968] give a concise statement of the general problem and a detailed discussion of a particular type of quality change.[1]

1. Preliminaries

An individual's tastes can be represented by a preference ordering defined over the commodity space. Let $X = (x_1,\ldots,x_n)$ and $X' = (x'_1,\ldots,x'_n)$ denote commodity bundles and write $X \; R \; X'$ for "X is at least as good as X'." The binary relation R is a preference

ordering. If R satisfies the usual conditions (completeness, reflexivity, transitivity, convexity, continuity), then there exists a real valued function, U(X), which "represents" the preference ordering in the sense that X R X′ if and only if $U(X) \geqq U(X')$. We call U(X) the "direct utility function" corresponding to the preference ordering R, and we sometimes write it as U(X;R) to emphasize the particular preference ordering, R, which the utility function represents. If U(X) is a direct utility function corresponding to the preference ordering, R, then any increasing monotonic transformation of U is also a direct utility function representing R.

The "ordinary demand functions" are found by maximizing the direct utility function subject to the budget constraint. Let $P = (p_1,\ldots,p_n)$ denote the price vector, and μ total expenditure.[2] We denote the ordinary demand functions by $x_i = h^i(P,\mu)$ or $x_i = h^i(P,\mu;R)$, when we need to indicate explicitly the preference ordering from which the demand functions are derived. In vector form, $X = H(P,\mu)$ or $X = H(P,\mu;R)$.

We say that a system of demand functions exhibits "expenditure proportionality" if $h^i(P,\mu) = \gamma^i(P)\mu$ for all i. This is equivalent to requiring all income elasticities to be 1 or, equivalently, all income-consumption curves to be rays from the origin. We say that a preference ordering is "homothetic to the origin" if it can be represented by a utility function which is an increasing monotonic transformation of a function homogeneous of degree 1: $U(X) = T[g(X)]$, $g(\lambda X) = \lambda g(X)$. We often call the utility function itself "homothetic". It is well known that a system of demand functions exhibits expenditure proportionality if and only if it is generated by a preference ordering which is homothetic to the origin.

The "compensated" or "constant utility" demand functions are found by minimizing the cost of attaining a particular indifference curve. We select a direct utility function to represent the preference ordering and denote the level of utility corresponding to the indifference curve by s; we denote the compensated demand functions by $x_i = f^i(P,s)$ or $x_i = f^i(P,s;R)$, or, in vector form, $X = F(P,s)$.[3] The role of s, the level of utility associated with the indifference curve, requires some explanation. The compensated demand functions depend on the particular indifference curve chosen but not on the particular utility function selected to represent the preference ordering. It would be more accurate to denote the compensated demand functions by $x_i = f^i(P,s;R,U)$ to indicate that the interpretation

of a particular numerical value of s depends on the direct utility function selected to represent R. Usually, this elaborate notation is unnecessary.

The direct utility function represents preferences by assigning a number to each X in the commodity space. One collection of goods is assigned a higher number than another if and only if the first is preferred to the second. In a similar manner we define an indirect utility function which represents preferences by assigning numbers to "price-expenditure situations", (P,μ): one price-expenditure situation is assigned a higher number than another if and only if the first is preferred to the second. The difficulty is that we have not yet defined what it means to say that one price-expenditure situation is preferred to another. The obvious meaning is that the best collection of goods available in the first situation is preferred to the best collection available in the second. Since $X = H(P,\mu;R)$ is the best collection of goods available at (P, μ), we say that (P, μ) is preferred to (P', μ') if and only if $X = H(P,\mu;R)$ is preferred to $X' = H(P',\mu';R)$. We can use the direct utility function, $U(X;R)$, to assign numbers to price-expenditure situations and thus to define the indirect utility function $\Psi(P,\mu;R)$:

$$\Psi(P,\mu;R) = U[H(P,\mu;R);R].$$

Thus, the indirect utility function is the maximum value of the direct utility function attainable in a particular price-expenditure situation:

$$\Psi(P,\mu;R) = \max U(X;R) \qquad \text{subject to} \qquad \Sigma\, p_k x_k \leq \mu.$$

The ordinary demand functions are related to the indirect utility function by

$$h^i(P,\mu) = -\frac{\dfrac{\partial \Psi(P,\mu)}{\partial p_i}}{\dfrac{\partial \Psi(P,\mu)}{\partial \mu}}$$

This result is often called Roy's theorem; since we use it repeatedly, we sketch a proof. (1) Differentiate $\Psi(P,\mu) = U[H(P,\mu)]$ with respect to p_i, replace U_k by $-\lambda p_k$ (from the first order conditions from maximization of the direct utility function), and replace $\Sigma\, p_k h_i^k$

by $-h^i$ (this follows from differentiating the budget constraint with respect to p_i). This yields $\Psi_i = \lambda h^i$. (2) Differentiate $\Psi(P,\mu) = U[H(P,\mu)]$ with respect to μ, replace U_k by $-\lambda p_k$, and recognize that $\Sigma\, p_k h^k_\mu = 1$. This yields $\Psi_\mu = -\lambda$. Hence, $-\Psi_i/\Psi_\mu = h^i$.

Finally, preferences can be represented by the "expenditure function", E(P,s), or E(P,s;R), analogous to the cost function in production theory.[4] In production theory the cost function shows the minimum cost of attaining a given level of output or, equivalently, a given isoquant; in consumer theory it shows the minimum expenditure required to attain a given level of utility or, strictly speaking, a given indifference curve. The expenditure function is related to the compensated demand functions by $E(P,s) = \Sigma\, p_k f^k(P,s)$.

Alternatively, the expenditure function may be derived from the indirect utility function by solving

$$s = \Psi(P,\mu)$$

for μ. Since

$$\frac{\partial E(P,s)}{\partial p_i} = -\frac{\dfrac{\partial \Psi[P,E(P,s)]}{\partial p_i}}{\dfrac{\partial \Psi[P,E(P,s)]}{\partial \mu}} = h^i[P,E(P,s)],$$

we have

$$\frac{\partial E(P,s)}{\partial p_i} = f^i(P,s).$$

2. Basic Theory

The cost-of-living index is the ratio of the minimum expenditures required to attain a particular indifference curve under two price regimes. We denote the cost-of-living index by $I(P^a,P^b,s,R)$:

$$I(P^a,P^b,s,R) = \frac{E(P^a,s,R)}{E(P^b,s,R)}$$

The notation emphasizes that the index depends not only on the two sets of prices, P^a and P^b, but also on an initial choice of an indifference map or preference ordering, R, and the choice of a base indifference curve, s, from that map. One set of prices is called "reference prices" and the other, "comparison prices". If comparison prices are twice the reference prices, the index is 2; if they are one-half the reference prices, the index is one-half. In our notation the comparison prices are the first n arguments of the index function, and reference prices, the next n arguments. Interchanging comparison and reference prices yields a new index which is the reciprocal of the original index:

$$I(P^a,P^b,s,R) = \frac{1}{I(P^b,P^a,s,R)} .$$

Either set of prices may be designated the reference set; a choice must be made, but it is a choice without substantive implications. Usually, we will denote comparison prices by P^a and reference prices by P^b. The reader is cautioned against calling the reference prices "base" prices; we reserve the term "base" to denote the indifference curve on which the index is predicated or the preference ordering to which it belongs.

Strictly speaking, the cost-of-living index depends only on the comparison prices, the reference prices, and the base indifference curve. It does not depend on the indifference map to which the base curve belongs. However, it is useful and realistic to imagine that the base indifference curve is selected by a two-stage procedure: first, a base map is chosen, and then a base curve is chosen from the map. Treating the base curve as part of a base map leads one to investigate the sensitivity of the index to the choice of the base curve from a particular map.

The logic of the cost-of-living index is best understood by interpreting it in the twin contexts of comparisons over time and comparisons over space. Features of the index which are easily overlooked in one context often stand out sharply in the other. This is particularly true of the role of the base preference ordering and base indifference curve.

For example, to construct a cost-of-living index to compare prices in Paris with those in Tokyo, we must specify the preference ordering on which the comparison is to be based. The Japanese government, considering how much to pay its diplomats in Paris, would presumably use Japanese tastes, while the French government would use French tastes. As is customary in discussing international price comparisons, we ignore differences in tastes within countries. But suppose the U.S. government wants to compare prices in Paris with those in Tokyo to decide on appropriate salary differentials for its diplomats. The comparison should be based on U.S. tastes. In principle, this is extremely important because it underscores the fact that the base preference ordering need not be one which is associated either with the reference prices or the comparison prices.

In intertemporal comparisons we often conceal the choice of a base preference ordering. To compare U.S. prices in 1970 with U.S. prices in 1969, it seems "obvious" and "natural" to use U.S. preferences. As a practical matter it seems likely that this would be appropriate in the majority of problems of this type although they would not be appropriate for the French government to use when deciding how much to increase the salary of its diplomats in Washington. The principal difficulty is that specifying "U.S. preferences" does not resolve the problem unless U.S. preferences are constant over time. Otherwise, it identifies a class of preference orderings from which the appropriate one still must be chosen. At first glance we have reduced the number of admissible preference orderings to two: U.S. preferences in 1969 and in 1970. But there is no reason why the comparison should not be based on U.S. preferences in 1971, 1958, or any other year. Fisher and Shell [1968] argue that current tastes provide a more appropriate indicator of the welfare effects of price changes than past tastes, but they are concerned only with the choice between 1969 and 1970 tastes and do not consider the possibility of basing the comparison on 1971 preferences.

The case of endogenous taste change is conceptually more difficult. If tastes change because of habit formation, as in Pollak [1970], then the appropriate base preference ordering may be a long-run pseudo preference ordering which generates the long-run demand functions rather than any particular short-run demand function with its implied dependence on the historic time path of consumption.[5] Endogenous taste change has received little systematic attention in economic theory; if tastes are endogenous, the validity of individual preferences as a touchstone of social welfare must be re-examined.

We conclude this section by enumerating the properties of the cost-of-living index which follow directly from its definition and the properties of the cost function:

$$I(P,P,s,R) = 1 \qquad \text{P1.}$$

That is, if the comparison prices are equal to the reference prices, the value of the index is 1.

$$I(\lambda P,P,s,R) = \lambda \qquad \text{P2.}$$

That is, if the comparison prices are proportional to the reference prices, then the value of the index is equal to the factor of proportionality.

$$I(P,\lambda P,s,R) = \frac{1}{\lambda} \qquad \text{P3.}$$

That is, if the reference prices are proportional to the comparison prices, then the value of the index is the reciprocal of the factor of proportionality.

$$I(\lambda P^a,\lambda P^b,s,R) = I(P^a,P^b,s,R). \qquad \text{P4.}$$

If the comparison prices and the reference prices are multiplied by a common factor, the value of the index is unchanged.

$$I(P^b,P^a,s,R) = 1/I(P^a,P^b,s,R). \qquad \text{P5.}$$

If the comparison and the reference prices are interchanged, then the new index is the reciprocal of the old.

$$\text{If } P^{a'} \geq P^a, \text{ then } I(P^{a'},P^b,s,R) \geq I(P^a,P^b,s,R). \qquad \text{P6.}$$

That is, if one set of comparison prices is higher than another, the index corresponding to the first is higher than that corresponding to the second. If the index is differentiable, we can express this monotonicity property as

$$\frac{\partial I}{\partial p_i^a}(P^a, P^b, s, R) \geq 0.$$

The strict inequality holds if all goods are consumed everywhere in a neighborhood of the initial price-expenditure situation. The property follows directly from the fact that an increase in any price cannot decrease the cost of attaining a particular indifference curve.

$$\min \left\{ \frac{p_i^a}{p_i^b} \right\} \leq I(P^a, P^b, s, R) \leq \max \left\{ \frac{p_i^a}{p_i^b} \right\}. \qquad \text{P7.}$$

The cost-of-living index for any base indifference curve lies between the smallest and the largest "price relative", p_i^a/p_i^b. To prove this we set μ^b so that $\Psi(P^b, \mu^b; R) = s$ and, therefore, $\mu^b = E(P^b, s; R)$. It suffices to show

$$\min \left\{ \frac{p_i^a}{p_i^b} \right\} \leq \frac{E(P^a, s; R)}{\mu^b} \leq \max \left\{ \frac{p_i^a}{p_i^b} \right\}$$

or, equivalently,

$$\mu^b \min \left\{ \frac{p_i^a}{p_i^b} \right\} \leq E(P^a, s; R) \leq \mu^b \max \left\{ \frac{p_i^a}{p_i^b} \right\}.$$

We now proceed with an overcompensation argument and an undercompensation argument. (1) If you give the individual $\mu^b \max \frac{p_i^a}{p_i^b}$, then he cannot be worse off than he was at P^b, μ^b because, regardless of what collection of goods he purchased at that price-expenditure situation, he can buy the same collection now with expenditure μ^b max $\frac{p_i^a}{p_i^b}$. In particular, this is true even if he consumed only one good, and that happened

to be the good which experienced the largest price increase. (2) If you give an individual $\mu^b \min \frac{p_i^a}{p_i^b}$, then he cannot be better off than at P^b, μ^b because the new feasible set lies entirely within the old one except where they coincide at the vertex corresponding to the good whose price has experienced the smallest increase. These upper and lower bounds are important because they do not depend on knowing anything about preferences except that they satisfy the usual regularity conditions, and because they do not depend on knowing the quantities consumed in any price-expenditure situation.[6]

3. Laspeyres and Paasche Indexes

The theory of the cost-of-living index provides no criterion for choosing either the base map or the base curve on which the index is predicated. The upper and lower bounds on the cost-of-living index expressed in P7.,

$$\min \left\{ \frac{p_i^a}{p_i^b} \right\} \leq I(P^a, P^b, s, R) \leq \max \left\{ \frac{p_i^a}{p_i^b} \right\},$$

represent the best that can be done without additional assumptions. In this section we examine the cost-of-living index corresponding to two "indifference map-indifference curve" combinations which stand out as "natural" or "obvious" ones on which to base the index, namely, those which correspond to the reference situation and the comparison situation.

Consider an individual with preference ordering, R^a, who, facing prices P^a with expenditure μ^a chooses the basket of goods X^a : $X^a = H(P^a, \mu^a; R^a)$. Similarly, $X^b = H(P^b, \mu^b; R^b)$. The most suggestive interpretation is in terms of place-to-place comparisons. Suppose P^a and P^b denote prices in Paris and Tokyo, so R^a and R^b denote French and Japanese preferences. There are two "natural" or "obvious" indifference curves which stand out as candidates on which to base a cost-of-living index. If we are to use French tastes, it seems "natural" (although certainly not necessary) to consider the indifference curve attained by a Frenchman facing prices P^a with income μ^a. We define s^a by $s^a = \Psi(P^a, \mu^a; R^a)$. If we are to use Japanese tastes as our norm, it seems "natural" to consider

the indifference curve attained by a Japanese facing prices P^b with expenditure μ^b: $s^b = \Psi(P^b,\mu^b;R^b)$. Thus, the two "natural" (s,R) combinations on which to base a living index are (s^a,R^a) and (s^b,R^b). In the case of intertemporal comparisons the situation is identical. However, if tastes do not change over time, then $R^a = R^b$, and the two "natural" base indifference curves belong to the same preference ordering.

We have identified two "natural" indexes, $I(P^a,P^b,s^a,R^a)$ and $I(P^a,P^b,s^b,R^b)$. There are other indexes which have some claim to being called "natural", and the primacy attributed to these two may reflect no more than the fact that we have interesting theorems about them. Two other "natural" indexes are $I(P^a,P^b,s^{a*},R^a)$ where $s^{a*} = \Psi(P^a,\mu^b;R^a)$ and $I(P^a,P^b,s^{b*},R^b)$ where $s^{b*} = \Psi(P^b,\mu^a;R^b)$. The first is based on the indifference curve which could be attained by an individual with the map of the comparison situation, facing comparison prices with the expenditure of the reference situation. The second is based on the curve attained by an individual with the map of the reference situation facing reference prices but with comparison expenditure.

The two indexes which we identified as natural are of special interest because we can establish better bounds for them than we could in the general case. To establish these bounds, we define a fixed weight index, $J(P^a,P^b,\theta)$,

$$J(P^a,P^b,\theta) = \frac{\Sigma\theta_k p_k^a}{\Sigma\theta_k p_k^b}$$

where $\theta = (\theta_1,\ldots,\theta_n)$. The fixed weight index is a ratio of weighted sums of prices, but we could, without loss of generality, divide through by $\Sigma\, \theta_k$ and interpret the index as a ratio of weighted averages of prices. We shall interpret the weights as quantities of the goods in a market basket, so the index is the ratio of the cost of that market basket at prices P^a to its cost at prices P^b.

A fixed weight price index provides little useful information unless the weights are carefully chosen. Two obvious choices of weights are X^a and X^b. The fixed weight index with weights equal to X^b is called a "Laspeyres" index:

$$J(P^a,P^b,X^b) = \frac{\Sigma\, x_k^b p_k^a}{\Sigma\, x_k^b p_k^b}.$$

That is, the Laspeyres index is a fixed weight index with weights associated with the reference prices, P^b. We often write the Laspeyres index in the form

$$J(P^a,P^b,X^b) = \Sigma\, w_k^b \left(\frac{p_k^a}{p_k^b}\right)$$

where $w_k^b = \dfrac{x_k^b p_k^b}{\mu^b}$. That is, the Laspeyres index is a weighted average of the "price relatives", p_i^a/p_i^b, where the weights are the expenditure weights of the reference situation. To show the equivalence of these two forms, we write

$$J(P^a,P^b,X^b) = \frac{\Sigma\, x_k^b p_k^a}{\mu^b} = \Sigma \frac{x_k^b}{\mu^b}\, p_k^a$$

$$= \Sigma \frac{x_k^b p_k^b}{\mu^b} \frac{p_k^a}{p_k^b} = \Sigma\, w_k^b \left(\frac{p_k^a}{p_k^b}\right).$$

We use the fact that X^b is the market basket purchased by an individual with preferences R^b with expenditure μ^b at prices P^b to establish an upper bound on $I(P^a,P^b,s^b,R^b)$.

Theorem:

$$\min \left\{ \frac{p_i^a}{p_i^b} \right\} \leq I(P^a, P^b, s^b, R^b) \leq J(P^a, P^b, X^b) \, .$$

The lower bound is the one asserted in P7., but the upper bound is an improvement since

$$J(P^a, P^b, X^b) = \Sigma \, w_k^b \left(\frac{p_k^a}{p_k^b} \right) \leq \max \left\{ \frac{p_k^a}{p_k^b} \right\} .$$

since the w's are non-negative numbers which sum to 1. That is, the Laspeyres index is an upper bound on the cost-of-living index based on the indifference curve attained in the reference situation.

Proof: To show that $J(P^a, P^b, X^b)$ is an upper bound on $I(P^a, P^b, s^b, R^b)$ we write the latter as

$$\frac{E(P^a, s^b, R^b)}{E(P^b, s^b, R^b)} = \frac{E(P^a, s^b, R^b)}{\mu^b}$$

and the former as

$$\frac{\Sigma \, x_k^b p_k^a}{\Sigma \, x_k^b p_k^b} = \frac{\Sigma \, x_k^b p_k^a}{\mu^b} .$$

It suffices to show that

$$E(P^a, s^b, R^b) \leq \Sigma \, x_k^b p_k^a \, .$$

But this follows directly from the fact that the minimum cost of attaining s^b at prices P^a cannot be greater than the cost of X^b.

The Paasche index is the fixed weight index with weights equal to the market basket purchased at the comparison prices, P^a , with expenditure μ^a:

$$J(P^a,P^b,X^a) = \frac{\Sigma\, x_k^a p_k^a}{\Sigma\, x_k^a p_k^b}\,.$$

It is the ratio of the cost of buying the market basket X^a at prices P^a to its cost at P^b. The Paasche index is a lower bound on $I(P^a,P^b,s^a,R^a)$, the cost-of-living index corresponding to (s^a,R^a).

Theorem:

$$J(P^a,P^b,X^a) \le I(P^a,P^b,s^a,R^a) \le \max \left\{\frac{p_i^a}{p_i^b}\right\}.$$

Proof: The upper bound is the one established in P7. To establish the lower bound, we must show that

$$\frac{\Sigma x_k^a p_k^a}{\Sigma\, x_k^a p_k^b} \le \frac{E(P^a,s^a,R^a)}{E(P^b,s^a,R^a)}\,.$$

Since the numerators are equal, it suffices to show

$$\frac{1}{\Sigma\, x_k^a p_k^b} \le \frac{1}{E(P^b,s^a,R^a)}$$

or, equivalently,

$$E(P^b,s^a,R^a) \le \Sigma\, x_k^a p_k^b\ .$$

But this follows immediately since the minimum expenditure required to attain (s^a,R^a)

at prices P^b cannot exceed the cost of X^a at these prices.

To summarize: the Laspeyres index is a fixed weight index with weights corresponding to the market basket purchased in the reference situation. It is an upper bound on the cost-of-living index corresponding to the preference ordering and indifference curve attained in the reference situation. The Paasche index is a fixed weight index with weights corresponding to the market basket purchased in the comparison situation. It is a lower bound on the cost-of-living index corresponding to the preference ordering and indifference curve attained in the comparison situation. It is not true that the cost-of-living index lies between the Paasche and Laspeyres indexes. Instead, we have a lower bound on one cost-of-living index and an upper bound on another.

4. When the Index is Equal to Its Bounds

In Section 3 we established two important bounding theorems for the cost-of-living index:

$$\min \left\{ \frac{p_i^a}{p_i^b} \right\} \leq I(P^a,P^b,s^b,R^b) \leq J(P^a,P^b,X^b)$$

and

$$J(P^a,P^b,X^a) \leq I(P^a,P^b,s^a,R^a) \leq \max \left\{ \frac{p_i^a}{p_i^b} \right\} .$$

In this section we investigate the preference orderings for which the cost-of-living index coincides with one or the other of its bounds.

4.1 When the cost-of-living index is equal to the Laspeyres or Paasche bounds

It is well known that if the preference ordering is represented by a "fixed coefficient"

direct utility function

$$U(X) = \min\left\{\frac{x_i}{a_i}\right\}$$

then the cost-of-living index $I(P^a,P^b,s^b,R)$ is equal to the Laspeyres index $J(P^a,P^b,X^b)$. To show this we make use of the fact that the expenditure minimizing quantity of good i for attaining a level of utility s^b is given by $x_i^b = a_i s^b$. Hence, the cost of attaining s^b at prices P^b is given by $\Sigma\, x_k^b p_k^b = \Sigma\, a_k s^b p_k^b$ while the minimum expenditure required to attain s^b at prices P^a is given by $\Sigma\, x_k^b p_k^a = \Sigma\, a_k s^b p_k^a$. Hence, the Laspeyres index $J(P^a,P^b,X^b)$ coincides with the cost-of-living index $I(P^a,P^b,s^b,R)$.

The homothetic fixed coefficient case is not the only one in which $I(P^a,P^b,s^b,R) = J(P^a,P^b,X^b)$. Any preference ordering which does not permit substitution along its indifference curves implies a cost-of-living index, $I(P^a,P^b,s^b,R)$, which coincides with the Laspeyres index, $J(P^a,P^b,X^b)$. There is no need for the indifference map to be homothetic.

Theorem: The cost-of-living index coincides with the appropriate Laspeyres and Paasche bounds if and only if the preference ordering can be represented by a direct utility function of the generalized fixed coefficient form

$$U(X) = \min\left\{g^i(x_i)\right\} \qquad g^{i\prime}(x^i) > 0.$$

Proof: If the preference ordering is of the generalized fixed coefficient form, it is easily verified that the cost-of-living index coincides with the appropriate Laspeyres and Paasche bounds.

If the cost-of-living index coincides with its Laspeyres bound for all s^b, then

$$\Psi(P^b,m) = \Psi[P^a,\Sigma\, h^k(P^b,m)p_k^a]$$

for all m. Differentiating with respect to p_i^b and m, we find the ordinary demand functions:

$$-h^i(P^b,m) = \frac{\Sigma p_k^a \dfrac{\partial h^k}{\partial p_i^b}}{\Sigma p_k^a \dfrac{\partial h^k}{\partial m}} .$$

Hence,

$$\Sigma p_k^a \left[\frac{\partial h^k}{\partial p_i^b} + h^i \frac{\partial h^k}{\partial m} \right] = 0$$

or, differentiating with respect to p_j^a,

$$\frac{\partial h^j}{\partial p_i^b} + h^i \frac{\partial h^j}{\partial m} = 0 .$$

This, of course, implies that the substitution effects are zero. We next show that the demand functions $h^2,\ldots,h^n$ can each be written as functions of h^1:

$$h^i(P,m) = \delta^i[h^1(P,m)] .$$

We do this by showing that the ratio of the partial derivatives is equal

$$\frac{\dfrac{\partial h^j}{\partial p_i}}{\dfrac{\partial h^j}{\partial m}} = \frac{\dfrac{\partial h^1}{\partial p_i}}{\dfrac{\partial h^1}{\partial m}} \quad \text{for all i,j} .$$

Equality of the partial derivatives follows from our characterization of the substitution

effect; indeed, the common value of the ratios is h^1. Substituting these demand functions into the direct utility function yields the indirect utility function

$$s = \Psi(P,\mu) = U[h^1(P,\mu),\delta^2(h^1),\ldots,\delta^n(h^1)] = \delta[h^1(P,\mu)] .$$

This implies $x_1 = f^1(P,s) = f^1(s)$, and, hence, $x_i = f^i(P,s) = f^i(s)$. This implies that the direct utility function is of the generalized fixed coefficient form where g^i is the inverse of f^i.

The Laspeyres (Paasche) index may coincide with the cost-of-living index to which it is the upper (lower) bound for a particular value of s, say s^*, but not for all s. This occurs if the indifference curve corresponding to s^* is of the fixed coefficient form. In another context Marjorie McElroy [1969] has provided an interesting example of such an indifference map; she constructed it by allowing the ''necessary basket'' of a linear expenditure system to coincide with the ''bliss point'' of an additive quadratic.[7] At the critical point, the Laspeyres and Paasche indexes coincide with the corresponding cost-of-living index.

4.2 When the cost-of-living index is equal to the ''other bounds''

In Section 4.1 we showed that the cost-of-living index coincides with the appropriate Laspeyres or Paasche bounds if and only if the preference ordering is of the generalized fixed coefficient form. We now examine the two forgotten bounds: $\left\{\frac{p_i^a}{p_i^b}\right\}$, the lower bound of $I(P^a,P^b,s^b,R^b)$ and max $\left\{\frac{p_i^a}{p_i^b}\right\}$, the upper bound of $I(P^a,P^b,s^a,R^a)$. For what preference orderings do these cost-of-living indexes coincide with their bounds?

If the preference ordering can be represented by a linear direct utility function

$$U(X) = \Sigma\, a_k x_k$$

then the indifference curves are parallel lines, and all goods are "perfect substitutes". The minimum cost of attaining the indifference curve s is given by

$$E(P,s,R) = \min \left\{ \frac{p_i}{a_i} s \right\} = s \min \left\{ \frac{p_i}{a_i} \right\}.$$

Hence, the cost-of-living index is

$$I(P^a,P^b,s,R) = \frac{\min \left\{ \frac{p_i^a}{a_i} \right\}}{\min \left\{ \frac{p_i^b}{a_i} \right\}}.$$

The ordinary demand functions corresponding to this utility are not single valued. If $\frac{p_1}{a_1} = \frac{p_2}{a_2} = \ldots = \frac{p_n}{a_n}$, then the budget line coincides with the indifference curve corresponding to $s = \frac{\mu a_i}{p_i}$, and the consumer is indifferent among all commodity bundles which exhaust his expenditure. For any other configuration of relative prices some goods will not be consumed.

Suppose that an individual's preferences are represented by a linear utility function and that when facing prices P^b with expenditure μ^b, he consumes all goods in positive quantities. Then

$$\frac{p_1^b}{a_1} = \frac{p_2^b}{a_2} = \ldots = \frac{p_n^b}{a_n} = r, \text{ so the cost of living index is given by}$$

$$I(P^a,P^b,s^b,R) = \frac{1}{r} \min \left\{ \frac{p_i^a}{a_i} \right\} = \min \left\{ \frac{p_i^a}{a_i} \right\} \frac{1}{r} = \min \left\{ \frac{p_i^a}{a_i} \frac{a_i}{p_i^b} \right\} = \min \left\{ \frac{p_i^a}{p_i^b} \right\}.$$

A similar result holds for the generalized linear direct utility function, $s = U(X)$, defined implicitly by

$$\Sigma\, \alpha^k(s)\, x_k = w(s),$$

where $\alpha^1,\ldots,\alpha^n$ and w are functions of s. The cost-of-living is given by

$$I(P^a,P^b,s,R) = \frac{\min\left\{\dfrac{p_i^a}{\alpha^i(s)}\right\}}{\min\left\{\dfrac{p_i^b}{\alpha^i(s)}\right\}}.$$

The indifference curves are linear, but they are not parallel; of course, they cannot intersect in the positive orthant, but there is no reason to rule out intersections of extensions of the indifference curves outside the commodity space. If all goods are consumed at (P^b,μ^b),

then it is easy to show that $I(P^a,P^b,s^b,R) = \min\left\{\dfrac{p_i^a}{p_i^b}\right\}$. This is what one would expect since the relevant characteristic in our previous example is that the indifference curves are linear; whether or not they are parallel is irrelevant.

We emphasize that it was necessary to assume that all goods are consumed at (P^b,μ^b). Suppose $p_i^a = p_i^b$, $i = 2,\ldots,n$ and $p_1^a < p_1^b$. If x_1 were not consumed at (P^b,μ^b) because it was too expensive, then a small decrease in its price, all other prices remaining constant, will not affect the cost-of-living index; x_1 will still be too expensive and will not be consumed. Hence, the value of the index would be 1, not $\min\left\{\dfrac{p_i^a}{p_i^b}\right\}$.

Theorem: The cost-of-living index coincides with the appropriate "other bounds", $\min\left\{\dfrac{p_i^a}{p_i^b}\right\}$ or $\max\left\{\dfrac{p_i^a}{p_i^b}\right\}$, if and only if the preference ordering can be represented by a

generalized linear utility function

$$\Sigma\, \alpha^k(s)\, x_k = w(s)\ ,$$

and all goods are consumed in positive quantities in the base situation.

Proof: We have already proved that if the utility function is of this form, then $I(P^a,P^b,s^b,R)$ coincides with $\min \left\{\dfrac{p_i^a}{p_i^b}\right\}$.

We now show that the preference ordering corresponding to the generalized linear direct utility function is the only one for which

$$I(P^a,P^b,s^b,R) = \min \left\{\frac{p_i^a}{p_i^b}\right\}$$

provided that all goods are consumed at prices P^b with expenditure μ^b. We cannot vary the p_i^b's because such variations may invalidate the hypothesis that all goods are consumed. We can, however, vary the p_i^a's. If $\dfrac{p_1^a}{p_1^b} < \dfrac{p_i^a}{p_i^b}$ for all $i \neq 1$, then in a neighborhood of P^a the index depends only on p_1^a and is independent of $p_2^a,\ldots,p_n^a$. Hence, in that neighborhood $E(P^a,s^b)$ depends only on p_1^a and s. Since the compensated demand functions are the derivatives of the expenditure function,

$$f^i(P^a,s^b) = \frac{\partial\, E(P^a,s^b)}{\partial\, p_i^a} = 0 \qquad i \neq 1.$$

That is, only good 1 is consumed. This is the case in every price-expenditure situation unless

there are "ties", and this implies linear indifference curves.

We now turn briefly to the index $I(P^a,P^b,s^a,R^a)$ and its upper bound $\max\left\{\frac{p_i^a}{p_i^b}\right\}$. If the preference ordering can be represented by the generalized linear direct utility function, then

$$I(P^a,P^b,s^a,R^a) = \frac{\min\left\{\frac{p_i^a}{\alpha_k(s^a)}\right\}}{\min\left\{\frac{p_i^b}{\alpha_k(s^a)}\right\}}.$$

If all goods are consumed at (P^a,μ^a), then

$$\frac{p_1^a}{\alpha_1(s^a)} = \frac{p_2^a}{\alpha_2(s^a)} = \ldots = \frac{p_n^a}{\alpha_n(s^a)} = r$$

so that the cost-of-living index becomes

$$I(P^a,P^b,s^a,R) = \frac{r}{\min\left\{\frac{p_i^b}{\alpha_i(s^a)}\right\}} = \frac{1}{\min\left\{\frac{p_i^b}{\alpha_i(s^a)}\,\frac{1}{r}\right\}} \doteq \frac{1}{\min\left\{\frac{p_i^b}{p_i^a}\right\}}$$

$$= \max\left\{\frac{p_i^a}{p_i^b}\right\}.$$

This result is precisely analogous to that obtained for $I(P^a,P^b,s^b,R)$, as indeed it must be.

The importance of the existence of preference orderings for which the cost-of-living index actually attains its "other bounds" lies in its immediate implication that these bounds are "best bounds". That is, if anyone claims to have found better bounds for the cost-of-living index, we can always find an admissible preference ordering whose cost-of-living

index lies outside the proposed bounds. Although our "other bounds" may not seem as satisfying or as useful as the Laspeyres and Paasche bounds, our demonstration that they correspond to the generalized linear utility function shows that it is not our lack of ingenuity but the inherent logic of the situation which prevents us from finding better ones.

5. The Base Indifference Curve

In this section I examine how the choice of the base indifference curve affects the cost-of-living index.

5.1 Expenditure proportionality and homothetic indifference maps

If the preference ordering is homothetic to the origin, then the implied cost-of-living index is independent of the particular indifference curve chosen as a base. That is, if R is homothetic to the origin, $I(P^a,P^b,s,R)$ is independent of s. To prove this, we use the fact that if a preference ordering is homothetic to the origin, then it can be represented by a direct utility function homogeneous of degree 1; the implied demand functions exhibit expenditure proportionality; and the indirect utility function can be written in the form

$$s = \Psi(P,\mu) = \phi(P)\mu$$

where $\phi(P)$ is homogeneous of degree -1. Hence, the expenditure function is given by

$$\mu = E(P,s) = \frac{s}{\phi(P)}$$

and the cost-of-living index by

$$I(P^a,P^b,s,R) = \frac{E(P^a,s,R)}{E(P^b,s,R)} = \frac{\phi(P^b)}{\phi(P^a)}$$

which is independent of s.

The converse of this result also holds: the cost-of-living index is independent of the base indifference curve if and only if the preference ordering is homothetic to the origin. Instead of showing this directly, we first introduce what appears to be a roundabout way of specifying the base indifference curve from a given preference ordering R. We specify the base indifference curve as the one corresponding to a base level of expenditure, m, at reference prices, P^b. The level of utility corresponding to the base indifference curve is given by $s = \Psi(P^b,m)$. It is often convenient to write the index as a function of m rather than s. We write $I^*(P^a,P^b,m,R) = I(P^a,P^b,\Psi(P^b,m),R)$. The index $I(P^a,P^b,s,R)$ is independent of s if and only if $I^*(P^a,P^b,m,R)$ is independent of m.

The specification of the index as a function of m rather than s is more than a useful mathematical trick. In practice, it is a sensible, convenient and commonly-used method of specifying the base indifference curve. That is, the base indifference curve is specified to be the indifference curve from a given base indifference map attainable by an individual with a particular expenditure at base period prices. In fact, we cannot interpret the index $I(P^a,P^b,s,R)$ without additional information which enables us to attach some meaning to the numerical value of s. To do this we need either the direct utility function, the indirect utility function, or the expenditure function. The cost-of-living index, $I^* = I^*(P^a,P^b,m,R)$, is defined implicitly by

$$\Psi(P^b,m) = \Psi(P^a,mI^*) .$$

Differentiating with respect to p_i^b and m and making use of the assumption that $\dfrac{\partial I^*(P^a,P^b,m,R)}{\partial m} = 0$ yields

$$\frac{\partial \Psi(P^b,m)}{\partial p_i^b} = \frac{\partial \Psi(P^b,mI^*)m}{\partial \mu} \frac{\partial I^*}{\partial p_i^b}$$

$$\frac{\partial \Psi(P^b,m)}{\partial m} = \frac{\partial \Psi(P^a,mI^*)I^*}{\partial \mu}$$

Hence,

$$h^i(P^b,m) = -\frac{\dfrac{\partial \Psi(P^b,m)}{\partial p_i^b}}{\dfrac{\partial \Psi(P^b,m)}{\partial m}} = -\left(\frac{1}{I^*}\ \frac{\partial I^*}{\partial p_i^b}\right) m \,.$$

Since the factor in parenthesis is independent of m, the demand functions exhibit expenditure proportionality, and, therefore, the preference ordering is homothetic to the origin.

We have just proved:

Theorem: The cost-of-living index is independent of the base indifference curve if and only if the preference ordering is homothetic to the origin.

This implies:

Theorem: If the preference ordering is homothetic to the origin, then

$$J(P^a,P^b,X^a) \leq I(P^a,P^b,s,R) \leq J(P^a,P^b,X^b) \,.$$

That is, if the preference ordering is homothetic, then the cost-of-living index lies between its Paasche and Laspeyres bounds. This follows immediately from our previous theorem which implies that the cost-of-living index is independent of the base indifference curve.

These results are important not because we believe that peoples' indifference maps are homothetic but because we believe they are not. Our theorem, therefore, implies that the cost-of-living index depends on the choice of the base level of expenditure. We now investigate the ways in which the preference ordering determines the relationship between the cost-of-living index and the base level of expenditure.

5.2 Demand functions locally linear in expenditure

We say that a system of demand functions is locally linear in expenditure if

$$h^i(P,\mu) = \chi_i(P) + \delta_i(P)\mu, \qquad i = 1,\ldots,n.$$

These demand functions are of substantially more empirical interest than those exhibiting expenditure proportionality. We now examine the form of the cost-of-living index implied by the preference ordering corresponding to these demand functions. W.M Gorman [1961] has shown that a system of demand functions is locally linear in expenditure if and only if its indirect utility function can be written in the form

$$s = \Psi(P,\mu) = \frac{\mu - f(P)}{g(P)},$$

where f(P) and g(P) are functions homogeneous of degree 1. The implied expenditure function is given by

$$\mu = E(P,s) \doteq f(P) + g(P)s$$

and the cost-of-living index by

$$I^*(P^a,P^b,m,R) = \frac{1}{m}\left[f(P^a) + g(P^a)\left(\frac{m - f(P^b)}{g(P^b)}\right)\right]$$

$$= \alpha(P^a,P^b) + \frac{1}{m}\beta(P^a,P^b)$$

where

$$\alpha(P^a,P^b) = \frac{g(P^a)}{g(P^b)}$$

$$\beta(P^a,P^b) = f(P^a) - f(P^b)\frac{g(P^a)}{g(P^b)}.$$

That is, if the demand functions are locally linear in expenditure, then the cost-of-living index is linear in the reciprocal of base expenditure. If f(P) = 0, then the indifference

map is homothetic to the origin, and the index is independent of m. As m approaches ∞, the cost-of-living index approaches a finite limit, and the influence of m becomes negligible. However, both of these assertions must be viewed cautiously since $f(P) = 0$ and $m \to \infty$ are inadmissible cases for certain g's. The quadratic case ($c = 2$) of Section 6.3 provides an illustration of both possibilities.

The converse of our characterization of the cost-of-living index also holds:

Theorem: The cost-of-living index depends linearly on the reciprocal of base expenditure if and only if the demand functions are locally linear in expenditure.

Proof: If $I = \alpha + \frac{\beta}{m}$, then the index is implicitly defined by

$$\Psi[P^b,m] = \Psi(P^a,mI) = \Psi(P^a,\alpha m + \beta) .$$

Differentiating with respect to p_i^b and m and recognizing that the demand functions are the negative of the ratios of these derivatives yields

$$h^i(P^b,m) = \frac{\alpha_i}{\alpha} m + \frac{\beta_i}{\alpha} .$$

5.3 Cost-of-living index linear in base expenditure

Theorem: If the cost-of-living index is linear in base expenditure, $I(P^a,P^b,m,R) = \alpha(P^a,P^b) + \beta(P^a,P^b)m$, then the demand functions are of the "Tornquist" form

$$h^i(P,\mu) = \frac{\alpha_i m + \beta_i m^2}{\alpha + 2\beta m} .$$

Proof: The index is implicitly defined by

$$\Psi[P^b,m] = \Psi(P^a,mI) = \Psi[P^a,\alpha m + \beta m^2] .$$

It is easily verified that the implied demand functions are of the Tornquist form.

This is an interesting result since Wold and Jureen [1953, p.3] suggest that demand functions of this form are not unreasonable. These results hold only for a limited range of values of m. The cost-of-living index cannot be linear for all m unless $\beta = 0$, in which case we are back to expenditure proportionality and homothetic indifference maps. If $\beta \neq 0$, the linear cost-of-living index would soon violate either the upper or lower bounds

$$\min \left\{ \frac{p_i^a}{p_i^b} \right\} \leq I^*(P^a, P^b, m, R) \leq \max \left\{ \frac{p_i^a}{p_i^b} \right\} .$$

6. Specific Preference Orderings

In this section I examine the cost-of-living-index formulae which correspond to specific preference orderings. In 6.1 I consider those corresponding to demand functions which exhibit expenditure proportionality and are generated by an additive direct utility function. In 6.2 I examine the cost-of-living indexes corresponding to demand functions which are locally linear in expenditure and are generated by an additive direct utility function. In 6.3 I consider the quadratic direct utility function. Irving Fisher's "ideal index", the geometric mean of the Laspeyres and the Paasche, is equal to the cost-of-living index if and only if the direct utility function is a homogeneous quadratic.

6.1 Additive utility functions and expenditure proportionality[8]

If an individual's utility function is additive and his demand functions exhibit expenditure proportionality, then his utility function belongs to the "Bergson family"

$$U(X) = \Sigma a_k \log x_k \qquad a_i > 0 \qquad \Sigma a_k = 1 \quad ^9 \tag{6.1.1}$$

$$U(X) = -\Sigma a_k x_k^c \qquad a_i > 0 \qquad c < 0 \tag{6.1.2}$$

$$U(X) = \Sigma a_k x_k^c \qquad a_i > 0 \qquad 0 < c < 1 \tag{6.1.3}$$

$$U(X) = \min \left\{ \frac{x_k}{a_k} \right\} \qquad a_i > 0 \tag{6.1.4}$$

The demand functions corresponding to (6.1.1) are of the form

$$h^i(P,\mu) = \frac{a_i \mu}{p_i} \tag{6.1.5}$$

while those corresponding to (6.1.2) and (6.1.3) are of the form

$$h^i(P,\mu) = \frac{\left(\dfrac{p_i}{a_i}\right)^{\frac{1}{c-1}} \mu}{\Sigma p_k \left(\dfrac{p_k}{a_k}\right)^{\frac{1}{c-1}}}. \tag{6.1.6}$$

The demand functions corresponding to the fixed coefficient case (6.1.4) are

$$h^i(P,\mu) = \frac{a_i \mu}{\Sigma a_k p_k}. \tag{6.1.7}$$

The indifference maps of these utility functions are identical with the isoquant maps of the C.E.S. production functions.

Since the demand functions exhibit expenditure proportionality, the corresponding indirect utility functions are of the form

$$\Psi(P,\mu) = \frac{\mu}{g(P)} \qquad (6.1.8)$$

where g(P) is homogeneous of degree 1. For the Cobb-Douglas case (6.1.1)

$$g(P) = \pi\, p_k^{a_k}. \qquad (6.1.9)$$

In the C.E.S. cases, (6.1.2) and (6.1.3),

$$g(P) = [\Sigma\, a_k^{-\frac{1}{c-1}} p_k^{\frac{1}{c-1}}]^{\frac{c-1}{c}} \qquad (6.1.10)$$

and in the fixed coefficient case (6.1.4)

$$g(P) = \Sigma\, a_k p_k\,. \qquad (6.1.11)$$

The indirect utility function (6.1.8) implies an expenditure function of the form

$$\mu = E(P,s) = g(P)s \qquad (6.1.12)$$

so the cost-of-living index is given by

$$I(P^a,P^b,s,R) = \frac{E(P^a,s,R)}{E(P^b,s,R)} = \frac{g(P^a)}{g(P^b)}. \qquad (6.1.13)$$

As we showed in Section 5, expenditure proportionality is a necessary and sufficient condition for independence of the base indifference curve.

Two cases deserve special mention.

Theorem: The cost-of-living index is a geometric mean of the price relatives with weight independent of s

$$I(P^a,P^b,s,R) = \pi\left(\frac{p_k^a}{p_k^b}\right)^{a_k} \qquad (6.1.14)$$

if and only if the utility function is of the Cobb-Douglas form (6.1.1).

Proof: It is easy to verify that if the utility function is of the form (6.1.1), then the cost-of-living index is given by (6.1.14), and the weights $(a_1,\ldots,a_n)$ are the budget shares: $a_i = p_i h^i(P,\mu)$. To prove the converse, write

$$\Psi(P^b,m) = \Psi\left[P^a, m\pi\left(\frac{p_k^a}{p_k^b}\right)^{a_k}\right]$$

differentiate with respect to p_i^b and m, and verify that the implied demand functions are given by $h^i(P^b,m) = a_i m/p_i^b$.

One way in which the Cobb-Douglas case is special is that the cost-of-living index is a function of the price relatives p_i^a/p_i^b. The class of preference orderings with cost-of-living indexes of this type is a generalization of the Cobb-Douglas class.

Theorem: If the cost-of-living index is a function of price relatives,

$$I(P^a,P^b,s,R) = \hat{I}\left(\frac{P^a}{P^b},s,R\right),$$

then the index is a geometric mean

$$I(P^a,P^b,s,R) = \pi\left(\frac{p_k^a}{p_k^b}\right)^{a^k(s)},$$

and the underlying preference ordering is a generalized Cobb-Douglas whose indirect utility function, $\Psi(P,\mu)$, is defined implicitly by

$$\Sigma\, \beta^k(s) \log p_k - \Sigma\, \beta^k(s) \log \mu = 1$$

where $a^i(s)$ is defined by $a^i(s) = \beta^i(s)/\Sigma\, \beta^k(s)$. [10]

where $a^i(s)$ is defined by $a^i(s) = \beta^i(s)/\Sigma\, \beta^k(s)$.[10]

It is sometimes thought that constructing a cost-of-living index is a matter of finding an appropriate way to combine the price relatives. This theorem shows that such a view is incorrect and that, except in the generalized Cobb-Douglas case, the comparison and reference prices do not enter the cost-of-living index in ratio form. Furthermore, the only admissible cost-of-living index based on price relatives is their geometric mean.

Proof: If the indirect utility function $\Psi(P,\mu)$ is implicitly defined by

$$\Sigma\, \beta^k(s) \log p_k - \log \mu\, \Sigma\, \beta^k(s) = 1 .$$

We solve for $\log \mu$ as a function of s:

$$\log \mu = \frac{\Sigma\, \beta^k(s) \log p_k - 1}{\Sigma\, \beta^k(s)}$$

The logarithm of the cost-of-living index is given by

$$\log I(P^a,P^b,s,R) = \log \frac{\mu^a}{\mu^b} = \log \mu^a - \log \mu^b =$$

$$= \frac{\Sigma\, \beta^k(s) \log p_k^a - \Sigma\, \beta^k(s) \log p_k^b}{\Sigma\, \beta^k(s)} .$$

Let $\alpha^i(s) = \beta^i(s)/\Sigma\, \beta^k(s)$. Then

$$\log I(P^a,P^b,s,R) = \Sigma\, a^k(s) \log \frac{p_k^a}{p_k^b} = \log \pi \left(\frac{p_k^a}{p_k^b}\right)^{a^k(s)}$$

so

$$I(P^a,P^b,s,R) = \pi\left(\frac{p_k^a}{p_k^b}\right)^{a^k(s)} .$$

The demand functions corresponding to this utility function are given by $h^i(P,\mu) = a^i(s)\mu/p_i$. This can be verified by differentiating

$$\Sigma\ \beta^k(s) \log p_k - \log \mu\ \Sigma\ \beta^k(s) = 1$$

with respect to p_i and and solving for $\dfrac{\partial \Psi}{\partial p_i}$ and $\dfrac{\partial \Psi}{\partial \mu}$.

We now show that this is the only preference ordering to yield a cost-of-living index which depends on price relatives. If

$$I(P^a,P^b,s,R) = \hat{I}\left(\frac{P^a}{P^b},s,R\right) ,$$

then

$$\frac{E(p_1^a,\ldots,\lambda p_i^a,\ldots,p_n^a,s)}{E(p_1^b,\ldots,\lambda p_i^b,\ldots,p_n^b,s)} = I(P^a,P^b,s,R) .$$

Differentiating with respect to λ and setting $\lambda = 1$ yields

$$p_i^a \frac{\partial E(P^a,s)}{\partial p_i^a} = I(P^a,P^b,s,R) p_i^b \frac{\partial E(P^b,s)}{\partial p_i^b}$$

so

$$\frac{f^i(P^a,s)p_i^a}{f^i(P^b,s)p_i^b} = I(P^a,P^b,s,R) = \frac{E(P^a,s)}{E(P^b,s)}$$

or, equivalently,

$$\frac{f^i(P^a,s)p_i^a}{E(P^a,s)} = \frac{f^i(P^b,s)p_i^b}{E(P^b,s)}.$$

That is, along an indifference curve the expenditure weight of each good is independent of prices: we denote the expenditure weight of good i by $a^i(s)$. But any system of demand functions of this form can be generated by the generalized Cobb-Douglas.

We showed in Section 4 that if the demand functions are generated by a homogeneous fixed coefficient utility function, then the Laspeyres and Paasche indexes coincide with each other, are independent of the base indifference curve, and coincide with the cost-of-living index. For completeness we restate that result here:

Theorem: If the direct utility function is of the homogeneous fixed coefficient form (6.1.4), then the cost-of-living index $I(P^a,P^b,s,R)$ coincides with the Laspeyres and Paasche indexes:

$$I(P^a,P^b,s,R) = \frac{\Sigma\, a_k p_k^a}{\Sigma\, a_k p_k^b} = \frac{\Sigma\, h^k(P,\mu)p_k^a}{\Sigma\, h^k(P,\mu)p_k^b} = J(P^a,P^b,X^b) = J(P^a,P^b,X^a) \qquad (6.1.15)$$

Since the Laspeyres index can be written as

$$J(P^a,P^b,X^b) = \Sigma\, w_k^b \left(\frac{p_k^a}{p_k^b}\right),$$

it might be thought that the homogeneous fixed coefficient case is one in which the cost-of-living index depends on price relatives. But this is not the case because the weights themselves depend on reference prices and not on the price relatives:

$$w_k^b = \frac{p_k^b h^k(P^b,\mu^b)}{\mu^b} = \frac{a_i p_i^b}{\Sigma\, a_k p_k^b}.$$

6.2 Additive utility functions and linear Engel curves

In this section I examine the cost-of-living index corresponding to demand functions which are locally linear in income and which are generated by additive direct utility functions. In Pollak [1971] I showed that the utility functions

$$U(X) = \Sigma\, a_k \log (x_k - b_k) \qquad a_i > 0,\ (x_i - b_i) > 0,\ \Sigma\, a_k = 1. \qquad (6.2.1)$$

$$U(X) = -\Sigma\, a_k(x_k - b_k)^c \qquad c < 0,\ (x_i - b_i) > 0. \qquad (6.2.2)$$

$$U(X) = \Sigma a_k(x_k - b_k)^{\,c} \qquad 0 < c < 1,\ a_i > 0,\ (x_i - b_i) > 0. \qquad (6.2.3)$$

$$U(X) = -\Sigma\, a_k\,(b_k - x_k)^c \qquad c > 1,\ a_i > 0,\ (b_i - x_i) > 0. \qquad (6.2.4)$$

$$U(X) = -\Sigma\, a_k^{\left(\frac{b_k - x_k}{a_k}\right)} \qquad a_i > 0. \qquad (6.2.5)$$

$$U(X) = \min\left\{\frac{x_k - b_k}{a_k}\right\} \qquad a_i > 0. \qquad (6.2.6)$$

The utility functions considered in Section 6.1 are special cases of (6.2.1), (6.2.2), (6.2.3) and (6.2.6), which correspond to these functions when all of the b's are 0.

The demand functions corresponding to (6.2.1) are of the form

$$h^i(P,\mu) = b_i - \frac{a_i}{p_i}\Sigma\, b_k p_k + \frac{a_i}{p_i}\mu\,. \tag{6.2.7}$$

This is the well-known Klein-Rubin [1947] linear expenditure system. The utility function (6.2.1) is a translated Cobb-Douglas. Similarly, (6.2.2), (6.2.3) and (6.2.6) are translations of the C.E.S. and fixed-coefficient cases considered in Section 6.1. The demand functions corresponding to (6.2.2), (6.2.3) and (6.2.4) are given by

$$h^i(P,\mu) = b_i - \gamma_i(P)\,\Sigma\, b_k p_k + \gamma_i(P)\mu \tag{6.2.8}$$

where

$$\gamma_i(P) = \frac{\left(\dfrac{p_i}{a_i}\right)^{\frac{1}{c-1}}}{\Sigma\, p_k\left(\dfrac{p_k}{a}\right)^{\frac{1}{c-1}}}\,.$$

The utility function corresponding to $c > 1$, (6.2.4), is not a generalization of an admissible C.E.S. case, but it includes the familiar additive quadratic ($c = 2$). The demand functions corresponding to (6.2.5) are given by

$$h^i(P,\mu) = b_i - \frac{a_i\,\Sigma\, p_k b_k}{\Sigma\, p_k a_k} + \frac{a_i\mu}{\Sigma p_k a_k} - a_i \log p_i + \frac{a_i \Sigma\, p_k a_k \log p_k}{\Sigma\, p_k a_k}\,.$$

The income-consumption curves are parallel straight lines. The demand functions corresponding to the translated fixed coefficient case, (6.2.6), are given by

$$h^i(P,\mu) = b_i - \left[\frac{a_i}{\Sigma\, a_k p_k}\right] b_k p_k + \left[\frac{a_i}{\Sigma\, a_k p_k}\right]\mu\,. \tag{6.2.10}$$

Since the demand functions are locally linear in expenditure, the indirect utility functions are of the form

$$\Psi(P,\mu) = \frac{\mu}{g(P)} - \frac{f(P)}{g(P)} \tag{6.2.11}$$

where f(P) and g(P) are homogeneous of degree 1. For the linear expenditure system, (6.2.1),

$$g(P) = \pi\, p_k^{a_k} \tag{6.2.12}$$

and

$$f(P) = \Sigma\, b_k p_k\ . \tag{6.2.13}$$

In the three C.E.S.-like cases, (6.2.2), (6.2.3) and (6.2.4),

$$g(P) = [\Sigma\, a_k^{\frac{-1}{c-1}} p_k^{\frac{1}{c-1}}]^{\frac{c-1}{c}} \tag{6.2.14}$$

and f(P) is given by (6.2.13). In the case of parallel income-consumption curves, (6.2.5),

$$g(P) = \Sigma\, a_k p_k \tag{6.2.15}$$

$$f(P) = (\Sigma\, a_k p_k)(\log \Sigma\, a_k p_k) + \Sigma\, b_k p_k - \Sigma\, a_k p_k \log p_k\ . \tag{6.2.16}$$

Finally, in the translated fixed-coefficient case, (6.2.6), f(P) is given by (6.2.13) and g(P) by (6.2.15).

Solving the indirect utility function, (6.2.11), for μ, we find that the expenditure function is of the form

$$\mu = E(P,s) = f(P) + g(P)s\ . \tag{6.2.17}$$

Hence, the cost of living index is of the form

$$I(P^a,P^b,s,R) = \frac{f(P^a) + g(P^a)s}{f(P^b) + g(P^b)s}. \tag{6.2.18}$$

It is often more useful to specify the base indifference curve in terms of base income m and to write the cost of living index in the form

$$I^*(P^a,P^b,m,R) = \alpha(P^a,P^b) + \frac{1}{m}\beta(P^a,P^b) \tag{6.2.19}$$

where

$$\alpha(P^a,P^b) = \frac{g(P^a)}{g(P^b)} \tag{6.2.20a}$$

$$\beta(P^a,P^b) = f(P^a) - f(P^b)\frac{g(P^a)}{g(P^b)}. \tag{6.2.20b}$$

Two cases are of special interest.

Theorem: The cost of living index corresponding to the linear expenditure system, (6.2.1), is of the form

$$I(P^a,P^b,s,R) = \frac{\Sigma\, b_k p_k^a + s\pi(p_k^a)^{a_k}}{\Sigma\, b_k p_k^b + s\pi(p_k^b)^{a_k}}. \tag{6.2.21}$$

This result was the conclusion of the article by Klein and Rubin [1947] in which they introduced the linear expenditure system. Interestingly enough, they entitled that paper "A Constant-Utility Index of the Cost of Living."

Theorem: If the direct utility function is of the translated fixed-coefficient form (6.2.6),

then the cost of living index $I(P^a,P^b,s^b,R)$ is equal to its Laspeyres upper bound $J(P^a,P^b,X^b)$; the cost of living index $I(P^a,P^b,s^a,R)$ is equal to its Paasche lower bound $J(P^a,P^b,X^a)$.

6.3 Quadratic direct utility function

The quadratic utility function is best treated in matrix form. Let B denote an n x 1 vector and A and n x n matrix. The direct quadratic utility function is given by

$$U(X) = X'B - \frac{1}{2}X'AX \tag{6.3.1}$$

where X is an n x 1 vector of commodities and primes denote transpose. We do not explicitly specify the regularity conditions for this utility function but restrict ourselves to a region of the commodity space in which it is well behaved. We also ignore the problems posed by goods which are not consumed; we work in a region of the price-expenditure space in which the set of goods consumed remains unchanged. As Wegge [1968] shows, this is not a trivial restriction in the case of the direct quadratic.

It is easy to verify that the demand functions corresponding to (6.3.1) are of the form

$$X = A^{-1}B + A^{-1}P\left[\frac{\mu - P'A^{-1}B}{P'A^{-1}P}\right], \tag{6.3.2}$$

where the expression in brackets is a scalar. Thus, the implied demand functions are locally linear in expenditure. The indirect utility function is given by

$$-\frac{1}{2}\frac{(\mu - P'A^{-1}B)^2}{P'A^{-1}P}.$$

If $P'A^{-1}P$ is positive within the admissible region of the price-expenditure space, we rewrite the indirect utility function in its "Gorman form":

$$\Psi(P,\mu) = \frac{\mu - P'A^{-1}B}{\sqrt{P'A^{-1}P}} \tag{6.3.3a}$$

If $P'A^{-1}P$ is negative, then

$$\Psi(P,\mu) = \frac{\mu - P'A^{-1}B}{\sqrt{-P'A^{-1}P}}. \tag{6.3.3b}$$

The corresponding expenditure function is given by

$$\mu = P'A^{-1}B + s\sqrt{\pm P'A^{-1}P}$$

and the cost of living index by

$$I(P^a,P^b,s,R) = \frac{P^{a\prime}A^{-1}B + s\sqrt{\pm P^{a\prime}A^{-1}P^a}}{P^{b\prime}A^{-1}B + s\sqrt{\pm P^{b\prime}A^{-1}P^b}} . \tag{6.3.4}$$

If B = 0, the direct utility function is a "homogeneous quadratic," and the cost-of-living index is independent of the base indifference curve:

$$I(P^a,P^b,s,R) = \sqrt{\frac{P^{a\prime}A^{-1}P^a}{P^{b\prime}A^{-1}P^b}} . \tag{6.3.5}$$

The homogeneous quadratic is of particular importance because it corresponds to Irving Fisher's Ideal Index, the geometric mean of a Laspeyres and a Paasche:

$$[J(P^a,P^b,H(P^b,\mu^b;R))J(P^a,P^b,H(P^a,\mu^a;R))]^{\frac{1}{2}} .$$

Furthermore, the homogeneous quadratic is the only preference ordering for which this is true:

Theorem: The cost of living index coincides with Fisher's Ideal Index

$$I^*(P^a,P^b,m,R) = [J(P^a,P^b,H(P^b,\mu^b;R))J(P^a,P^b,H(P^a,\mu^a;R))]^{\frac{1}{2}}$$

if and only if the preference ordering is a homogeneous quadratic.[11]

Proof: If the direct utility function is a homogeneous quadratic, then the Laspeyres and Paasche indexes are given by

$$J[P^a,P^b,H(P^b,\mu^b;R)] = \frac{P^{a\prime}A^{-1}P^b}{P^{b\prime}A^{-1}P^b}$$

$$J[P^a,P^b,H(P^a,\mu^a;R)] = \frac{P^{a\prime}A^{-1}P^a}{P^{b\prime}A^{-1}P^a}\ .$$

Since $P^{a\prime}A^{-1}P^b = P^{b\prime}A^{-1}P^a$, the ideal index is (6.3.5).

If the cost of living index coincides with the ideal index, then the Laspeyres index is independent of μ^b, and hence the demand functions exhibit expenditure proportionality. We write the indirect function in its Gorman form, $\Psi(P,\mu) = \mu/\phi(P)$, where $\phi(P)$ is a function homogeneous of degree 1. The demand functions are given by $h^i(P,\mu) = \phi_i(P)\mu/\phi(P)$. The cost of living index I is implicitly defined by

$$\Psi(P^b,m) = \Psi(P^a,mI)$$

so

$$\frac{\phi(P^a)}{\phi(P^b)} = I\ .$$

The product of the Laspeyres and Paasche indexes can be written as

$$\frac{\Sigma\, h^k(P^b,\mu^b)p_k^a}{\Sigma\, h^k(P^b,\mu^b)p_k^b}\;\frac{\Sigma\, h^k(P^a,\mu^a)p_k^a}{\Sigma\, h^k(P^a,\mu^a)p_k^b} = \frac{\Sigma\, \phi_k(P^b)p_k^a}{\Sigma\, \phi_k(P^b)p_k^b}\;\frac{\Sigma\, \phi_k(P^a)p_k^a}{\Sigma\, \phi_k(P^a)p_k^b}$$

$$= \frac{\Sigma\, \phi_k(P^b)p_k^a}{\phi(P^b)}\;\frac{\phi(P^a)}{\Sigma\, \phi_k(P^a)p_k^b}.$$

If the cost of living index is equal to the ideal index, then

$$\left(\frac{\phi(P^a)}{\phi(P^b)}\right)^2 = \frac{\Sigma\, \phi_k(P^b)p_k^a}{\phi(P^b)}\;\frac{\phi(P^a)}{\Sigma\, \phi_k(P^a)p_k^b}$$

or, equivalently,

$$\phi(P^a)\, \Sigma\, \phi_k(P^a)p_k^b = \phi(P^b)\, \Sigma\, \phi_k(P^b)p_k^a\,.$$

Differentiating with respect to p_i^a yields

$$\phi_i(P^a)\, \Sigma\, \phi_k(P^a)p_k^b + \phi(P^a)\, \Sigma\, \phi_{ki}(P^a)p_k^b = \phi(P^b)\phi_i(P^b)\,.$$

Differentiating this with respect to p_j^b yields

$$\phi_i(P^a)\phi_j(P^a) + \phi(P^a)\phi_{ji}(P^a) = \phi_j(P^b)\phi_i(P^b) + \phi(P^b)\phi_{ij}(P^b)\,.$$

Since the right hand side and the left hand side are equal regardless of the values of P^a and P^b, both must be independent of P and hence constant:

$$\phi_i(P)\phi_j(P) + \phi(P)\phi_{ji}(P) = c_{ij}\,.$$

Multiplying by p_j and summing over j yields

$$\phi_i(P)\phi(P) = \sum_j p_j c_{ij}$$

since the homogeneity of ϕ implies

$$\sum_j p_j \phi_{ji}(P) = 0 .$$

Multiplying by p_i and summing over i yields

$$[\phi(P)]^2 = \sum_i \sum_j p_i p_j c_{ij}$$

so

$$\phi(P) = \pm \sum_i \sum_j p_i p_j c_{ij}$$

which is the homogeneous quadratic.

We remark that if $c_{ij} = a_i a_j$, then

$$\phi(P) = \Sigma\, a_k p_k$$

and that

$$\Psi(P,\mu) = \frac{\mu}{\Sigma\, a_k p_k}$$

is the indirect utility function corresponding to the homogeneous fixed-coefficient utility function. It should be no surprise that the homogeneous fixed-coefficient case appears here, for this is the case in which the cost of living index coincides with both the Laspeyres and the Paasche indexes, and, hence, it must coincide with their geometric mean.

7. A Preference Field Quantity Index

As we have seen, the cost of living index provides a precise answer to a narrow and specific question. If one wishes to compare expenditures required to attain a particular base indifference curve at two sets of prices, then, by definition, the cost of living index is the appropriate index. But price indexes are often used to deflate an index of total expenditure to obtain an index of quantity or "real consumption." With less logic but the same purpose they are used to deflate indexes of money income to obtain indexes of real income or money wages to obtain real wages. Although the last two cases are clouded by problems involving saving and the labor-leisure choice, the purposes of these indexes is the measure "quantity."[12] In this section we show how a quantity index can be constructed by a procedure analogous to that used to construct the cost of living index. We call such an index a "preference field quantity index" to distinguish it from other types of quantity indexes and to suggest its relation to the "preference field price index" or cost of living index of Section 2. Before constructing the preference field quantity index, we give a careful summary of the logic which lies behind the preference field price index. In Section 8 we examine the conditions under which the preference field quantity index coincides with the quantity index obtained by using the cost of living index to deflate an index of expenditure.

7.1 The preference field price index

Given a preference ordering R and a base indifference curve s we defined the cost of living index by

$$I(P^a,P^b,s,R) = \frac{c_a}{c_b} \qquad (7.1.1)$$

where c_a and c_b are implicitly defined by

$$s = \Psi(P^a,c_a;R)$$

$$s = \Psi(P^b,c_b;R) .$$

In theoretical work involving indirect utility functions, it is standard practice to work with "normalized prices." We define y_i, the normalized price of the ith good, by $y_i = p_i/\mu$ and let Y denote the corresponding vector. The ordinary demand functions $h^i(P,\mu;R)$ can be written as $g^i(Y;R)$ since they are homogeneous of degree 0 in all prices and expenditure:

$$g^i(Y;R) = h^i(Y,1;R) = h^i\left(\frac{p_1}{\mu},\ldots,\frac{p_n}{\mu},1;R\right) = h^i(P,\mu;R)\,.$$

Similarly, the indirect utility function $\Psi(P,\mu;R)$ can be written as $\phi(Y;R)$ since it too is homogeneous of degree 0 in all prices and expenditure:

$$\phi(Y;R) = \Psi(Y,1;R) = \Psi\left(\frac{p_1}{\mu},\ldots,\frac{p_n}{\mu},1;R\right) = \Psi(P,\mu;R)\,.$$ [13]

The ordinary demand functions are related to the normalized indirect utility function by

$$g^i(Y) = h^i(Y,1) = \frac{\phi_i(Y)}{\Sigma\, y_k \phi_k(Y)}\,.$$

Using this new notation, the cost of living index is given by (7.1.1) where c_a and c_b are implicitly defined by

$$s = \phi\left(\frac{1}{c_a}P^a;R\right)$$

$$s = \phi\left(\frac{1}{c_b}P^b;R\right)\,.$$

We can give a straightforward interpretation of the cost of living index in terms of the indifference curves corresponding to the indirect utility function $\phi(Y)$. We remind the reader that in Figure 1 utility *increases* as you move *toward* the origin. We begin with the base preferences ordering (represented by the indifference map) and, from the map, the base indifference curve, which we denote by s. We let $Y^a = P^a$ and $Y^b = P^b$ denote the comparison and reference prices, respectively. In general, neither the reference nor the com-

parison prices will lie on the base indifference curve. We let $Y^{a'}$ denote the point at which the ray from the origin through Y^a intersects the base indifference curve. Similarly, $Y^{b'}$ denotes the intersection of the base indifference curve with the ray from the origin through Y^b.

It is instructive to decompose the construction of the cost of living index comparing Y^a with Y^b using the base indifference curve s into three separate parts: a comparison of Y^a with $Y^{a'}$ with $Y^{b'}$, and a comparison of $Y^{b'}$ with Y^b. First, we compare Y^a with $Y^{a'}$. That is, suppose the comparison prices Y^a and the reference prices (for our present purpose, $Y^{a'}$) lie on the same ray: $Y^a = c_a Y^{a'}$. Then the cost of living index $I(Y^a,Y^{a'},s,R)$ is equal to c_a, a result which is intuitively obvious and natural. Formal justification is provided by

$$s = \phi\left(\frac{1}{c_a}Y^a;R\right) = \phi(Y^{a'};R) .$$

If, for example, $Y^a = 4Y^{a'}$, and $Y^{a'}$ is on the base indifference curve, then $I(Y^a,Y^{a'},s,R) = 4$.

We now compare $Y^{a'}$ with $Y^{b'}$. That is, suppose the comparison prices are $Y^{a'}$ and the reference prices $Y^{b'}$. Since, by construction, both $Y^{a'}$ and $Y^{b'}$ lie on the base indifference curve, the value of index is 1: $I(Y^{a'},Y^{b'},s,R) = 1$.

Finally, we compare $Y^{b'}$ with Y^b. In this case the comparison prices $Y^{b'}$ and the reference prices Y^b lie on the same ray: $Y^b = c_b Y^{b'}$. Hence, the cost-of-living index $I(Y^{b'},Y^b,s,R)$ is equal to $1/c_b$, as common sense requires. For example, if $Y^b = 3Y^{b'}$, then $I(Y^{b'},Y^b,s,R) = \frac{1}{3}$

We must now combine these three results to obtain the cost-of-living index we originally sought: $I(Y^a,Y^b,s,R)$. It is at this stage that our argument loses some of its intuitive appeal. It is perfectly clear how we ought to compare two-price situations which lie on a common ray or two-price situations which lie on the base indifference curve, but it is less clear how we ought to proceed in other cases. Formally, we can proceed by introducing

a formal rule which enables us to combine indexes with a common base indifference curve:

$$I(Y^a,Y^c,s,R) = I(Y^a,Y^b,s,R)I(Y^b,Y^c,s,R) .$$

We saw in Section 2 that cost-of-living indexes can be combined in this way. But the intuitive justification for this step is not a completely comfortable one. To understand the way in which this multiplication rule operates, suppose that the reference prices lie on the base indifference curve but not on the same ray as the comparison prices. The index is then equal to the ratio of the comparison prices to the point where the comparison ray intersects the base indifference curve. In effect, we construct the index as if the reference prices were at this intersection. Our rationale for this is that the reference prices and the intersection lie on the base indifference curve, and we regard points on the base indifference curve as "equivalent".

We now take account of the fact that the reference prices Y^b need not lie on the base indifference curve. We do this by making use of $P^{b'}$, the point at which the reference ray intersects the base indifference curve. As we saw, if $Y^b = c_b Y^{b'}$, then $I(Y^{b'},Y^b,s,R) = 1/c_b$. Furthermore, if the comparison prices lie anywhere on the base indifference curve, the value of the index would still be $1/c_b$. If neither the comparison nor the reference prices lie on the base indifference curve, we construct the index by reducing both reference and comparison prices to the base indifference curve and making use of the "equivalence" of all price situations on the base indifference curve. Thus, in our example, the value of the index would be

$$I(Y^a,Y^b,s,R) = I(Y^a,Y^{a'},s,R)I(Y^{a'},Y^{b'},s,R)I(Y^{b'},Y^b,s,R) = 4 \times 1 \times \frac{1}{3} = \frac{4}{3} .$$

The procedure for constructing a cost-of-living index based on a particular indifference curve s can be summarized in three simple axioms:

$$\text{If} \quad s = \phi(Y^a;R) \quad \text{and} \quad s = \phi(Y^b;R), \quad \text{then } I(Y^a,Y^b,s,R) = 1 . \qquad \text{A1.}$$

ure 1

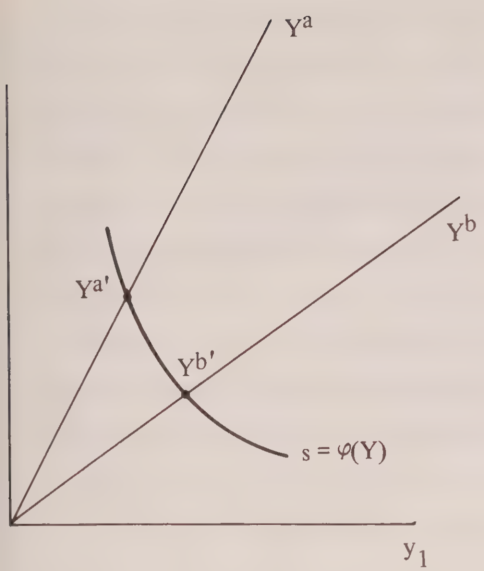
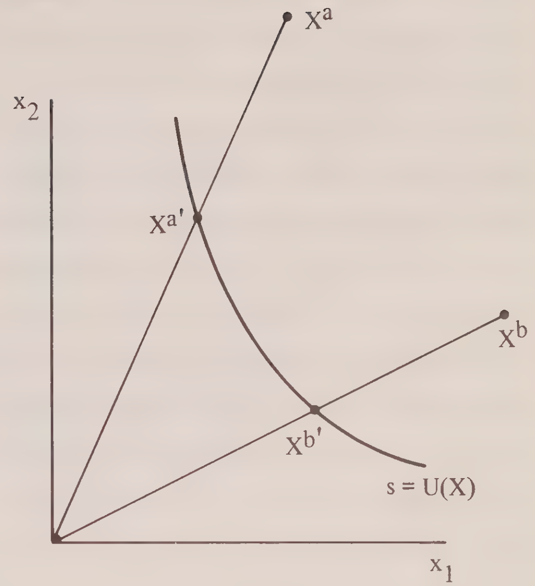


Figure 2



That is, if the comparison and the reference prices both lie on the base indifference curve, then the value of the index is 1.

$$\text{If} \quad s = \phi(Y^b;R) \quad \text{and} \quad Y^a = \lambda Y^b, \quad \text{then } I(Y^a,Y^b,s,R) = \lambda \,. \qquad \text{A2.}$$

In words, if the reference prices lie on the base indifference curve and the comparison prices lie on the same ray as the reference prices, then the value of the index is the factor of proportionality relating the comparison to the reference prices.

$$I(Y^a,Y^b,s,R)I(Y^b,Y^c,s,R) = I(Y^a,Y^c,s,R) \,. \qquad \text{A3.}$$

If two cost-of-living indexes are based on the same base indifference curve, and the reference prices in the first are the same as the comparison prices in the second, then the product of these two indexes is the cost-of-living index whose comparison prices are those of the first and whose reference prices are those of the second. Irving Fisher [1922, p.270] called this the "circular" property.

This axiomatic treatment of the procedure for constructing the cost-of-living index makes it clear that the base indifference map plays no role and that the base curve does all the

work. It is sometimes useful to discard the original indifference map and work only with the base indifference curve. Sometimes, however, it is useful to go a step further and think of the index as being constructed from a pseudo indifference map which is defined as the radial or homothetic blowup of the base indifference curve. Unless the original indifference map was homothetic to the origin, the pseudo map does not coincide with the original map. One advantage of introducing the pseudo map is that because it is homothetic to the origin, the cost-of-living index is independent of which pseudo indifference curve is treated as the base. In particular, without loss of generality, we may use the curve on which the reference prices lie as the base curve. On the pseudo map (but not on the original map unless it was homothetic) if the reference prices and the comparison prices lie on the same curve, then the value of the index is 1. Notice, however, that if the reference and comparison prices lie on the same indifference curve on the original map, and their common curve is not the base curve, then the value of the index need not be 1.

7.2 Preference field quantity indexes

We now use the same formal procedure to define a preference field quantity index as we used to define a preference field price index or cost of living index. That is, we define the preference field quantity index $Q(X^a,X^b,s,R)$ by

$$Q(X^a,X^b,s,R) = \frac{\theta_a}{\theta_b}$$

where θ_a and θ_b are defined by

$$s = U\left(\frac{1}{\theta_a}X^a\right)$$

and

$$s = U\left(\frac{1}{\theta_b}X^b\right).$$

The verbal and graphical interpretation of the preference field quantity index is essentially the same as that of the preference field price index. We begin by choosing a base

preference ordering and from it a base indifference curve. In general, neither the comparison quantities X^a nor the reference quantities X^b lie on the base indifference curve. Instead of comparing X^a directly with X^b, we compare X^a with a collection of goods $X^{a\prime}$ which lies on s and which is proportional to X^a. Similarly, we compare X^b with a collection $X^{b\prime}$ which lies on s and which is proportional to X^b. Graphically, the preference field quantity index can be represented by a diagram, Figure 2, identifical to that used to illustrate the construction of the cost of living index.

Since the preference field quantity index is less familiar than the cost of living index, it is useful to examine the meaning of this index. Formally, the axioms satified by this index are identical with those satisfied by the preference field price index:

$$\text{If } s = U(X^a) \text{ and } s = U(X^b) \text{ then } Q(X^a,X^b,s,R) = 1 \qquad \text{A1.}$$

$$\text{If } s = U(X^b) \text{ and } X^a = \lambda X^b \text{ then } Q(X^a,X^b,s,R) = \lambda \qquad \text{A2.}$$

$$Q(X^a,X^b,s,R)Q(X^b,X^c,s,R) = Q(X^a,X^c,s,R)\ . \qquad \text{A3.}$$

That is, (A1.) if both the comparison and reference quantities lie on the base indifference curve, then the value of the index is 1. (A2.) If the reference quantities lie on the base indifference curve, and the comparison quantities and the reference quantities lie on a common ray, then the value of the index is the ratio of the comparison to the reference quantities. (A3.) If two indexes have the same base, and the reference quantities of the first are equal to the comparison quantities of the second, then the product of the two indexes is an index whose comparison quantities are equal to those of the first and whose reference quantities are those of the second. These three axioms uniquely determine the preference field quantity index corresponding to any base indifference curve.

As with the preference field price index, the construction of the index depends only on the base indifference curve but not on the underlying preference ordering or indifference map. It is useful to consider the index in relation to a pseudo indifference map constructed as the homothetic blowup of the base indifference curve. On the pseudo indifference map the index is independent of the choice of the pseudo indifference curve used as the base. Thus, we can choose the curve corresponding to the reference quantities with no loss of

generality. If the comparison and reference quantities lie on the same pseudo indifference curve, then the value of the index is 1. But if the comparison and reference quantities lie on the same indifference curve on the original map, the value of the index need not be unity unless they both lie on the base indifference curve. This is important because it underscores the fact that the preference field quantity index does not compare the levels of utility corresponding to the reference and comparison quantities.

8. Price, Quantity, and Expenditure Indexes

Since expenditure is a scalar, an expenditure index $M(\mu_a,\mu_b)$ is naturally defined by $M(\mu_a,\mu_b) = \dfrac{\mu_a}{\mu_b}$. Such an index is independent of preferences. In this section we examine the relationships among price, quantity, and expenditure indexes. In particular, we are interested in the conditions under which the expenditure index $M(\mu^a,\mu^b)$ can be decomposed into the product of a price index and a quantity index or, equivalently, the conditions under which the quantity index is equal to the expenditure index deflated by a price index or, equivalently, the conditions under which the price index is equal to the expenditure index deflated by a quantity index. The answers to these questions depend on the types of price and quantity indexes which are admissible.

One approach is to require the price index to be a cost of living index and to define the quantity index by

$$Q(P^a,P^b,\mu^a,\mu^b,s,R) = \frac{M(\mu^a,\mu^b)}{I(P^a,P^b,s,R)}.$$

Alternatively, if we insist that the quantity index be a preference field index, we can define the price index by

$$I(X^a,X^b,\mu^a,\mu^b,s,R) = \frac{M(\mu^a,\mu^b)}{Q(X^a,X^b,s,R)}.$$

The drawback of this approach is that the derived indexes do not satisfy the axioms discussed in Section 7 unless the base preference ordering is homothetic. Furthermore, it is easy to verify that

$$Q(P^a,P^b,\mu^a,\mu^b,s,R)Q(P^b,P^c,\mu^b,\mu^c,s,R) = Q(P^a,P^c,\mu^a,\mu^c,s,R)$$

so that A3., the least intuitively appealing of our axioms, is always satisfied by a quantity index defined in this way. Hence, failure of the quantity index to satisfy our axioms implies a violation of either A1., which requires that if both the comparison and reference quantities lie on the base indifference curve, then the value of the index is 1, or A2., which requires that if the reference quantities lie on the base indifference curve and both the comparison and reference quantities lie on a common ray, then the index is the ratio of the comparison to the reference quantities. Both of these axioms have such strong intuitive appeal that a quantity index which violates them has little conceptual appeal. We show in this section that the preference field quantity index is equal to the expenditure index deflated by the cost of living index if and only if the preference ordering is homothetic to the origin. This implies that unless the preference ordering is homothetic, $Q(P^a,P^b,\mu^a,\mu^b,S,R)$ violates either A1. or A2.

We now require both the price and the quantity indexes to be preference field indexes and determine the class of preference orderings for which their product is equal to the expenditure index.

Theorem: $M(\mu^a,\mu^b) = I(P^a,P^b,s,R)Q[H(P^a,\mu^a;R),H(P^b,\mu^b;R),s,R]$, for all P^a,P^b,μ^a,μ^b, if and only if the preference ordering is homothetic to the origin.

The statement of the theorem requires that the result hold for only one base indifference curve, s, but the conclusion implies that if it holds for one base curve, then the preference ordering is such that it holds for every base curve.

Proof: If the preference ordering is homothetic to the origin, and the demand functions were generated by this preference ordering, we can write the direct utility function U(X) as a function homogeneous of degree 1; the indirect utility function $\Psi(P,\mu) = U[H(P,\mu)]$

$= U[H(P,1)]\mu$ can be written as $\dfrac{\mu}{\phi(P)}$ where $\phi(P)$ is homogeneous of degree 1, $\phi(P) = 1/U[H(P,1)]$.

The cost of living index is given by

$$I(P^a,P^b,s,R) = \frac{\phi(P^a)}{\phi(P^b)}.$$

The preference field quantity index is given by

$$Q(X^a,X^b,s,R) = \frac{\theta_a}{\theta_b} = \frac{U(X^a)}{U(X^b)}$$

since

$$s = U\left(\frac{1}{\theta_a}X^a\right) = \frac{1}{\theta_a}U(X^a)$$

and

$$s = U\left(\frac{1}{\theta_b}X^b\right) = \frac{1}{\theta_b}U(X^b)\,.$$

But $U(X^a) = \dfrac{\mu^a}{\phi(P^a)}$ and $U(X^b) = \dfrac{\mu^b}{\phi(P^b)}$, so the product of the preference field price and quantity indexes is μ^a/μ^b.

We prove the second part of the theorem by showing that if the expenditure index is equal to the product of the preference field price and quantity indexes, then the demand functions exhibit expenditure proportionality. We first observe that the quantity index must be homogeneous of degree 1 in μ^a since μ^a appears only in $M(\mu^a,\mu^b)$ and in $H(P^a,\mu^a)$. Let P^a be any price vector; set P^b equal to P^a. Choose μ^b so that $\Psi(P^b,\mu^b) = s$. Then $X^b = H(P^b,\mu^b)$ lies on the base indifference curve. If we set μ^a equal to μ^b, then $X^a =$

X^b and $Q(X^a,X^b,s,R) = 1$. Suppose $\mu^a = \lambda\mu^b$. Since the quantity index is homogeneous of degree 1 in λ, its value is λ. This means that X^a lies on the pseudo indifference curve which is a radial blowup of the base indifference curve by the scale factor λ. But X^a must lie in the feasible set defined by P^a,μ^a. This feasible set is a radial blowup of the feasible set defined by P^b,μ^b. If the indifference curves are strictly convex, X^b is the only point on the base indifference curve in the feasible set of P^b,μ^b. Hence, λX^b is the only point on the radial blowup of the base indifference curve which lies in the feasible set of P^a,μ^a. Hence, $X^a = \lambda X^b$ and the demand functions exhibit expenditure proportionality.

It might be argued that this theorem is not surprising since we have not allowed the base indifference curve to vary with μ^a or μ^b. If the base indifference curve were always chosen so that $s = \Psi(P^b,\mu^b)$, perhaps some non-homothetic preference ordering would suffice. We now show that this is not the case.

Theorem: $M(\mu^a,\mu^b) = I[P^a,P^b,\Psi(P^b,\mu^b;R);R]Q[H(P^a,\mu^a;R),H(P^b,\mu^b;R),\Psi(P^b,\mu^b;R),R]$, for all P^a,P^b,μ^a,μ^b if and only if the preference ordering is homothetic to the origin.

That is, allowing the base of the index to vary with μ^b does not permit any generalization of our theorem; clearly, allowing it to vary with μ^a would not.

Proof: The proof of the previous theorem showed that if the indifference map is homothetic, both the price and quantity indexes are independent of the base indifference curve, and their product is equal to the expenditure index.

To prove that only homothetic preference orderings will work, we first observe that the quantity index is homogeneous of degree 1 in μ^a. We proceed as in the proof of the previous theorem except that we are free to choose any initial value for μ^b.

It might still be objected (by analogy with the fact that the product of a Laspeyres price index and a Paasche quantity index is equal to the expenditure index) that we should not require the same base indifference curve to serve for both indexes. We now show that basing the price index on the reference situation and the quantity index on the comparison situation does not permit any generalization of our result.

Theorem:
$M(\mu^a,\mu^b) = I[P^a,P^b,\Psi(P^b,\mu^b;R);R]Q[H(P^a,\mu^a;R),H(P^b,\mu^b;R),\Psi(P^a,\mu^a;R),R]$ for all P^a,P^b,μ^a,μ^b if and only if the preference ordering is homothetic to the origin.

Proof: We have already seen that if the indifference map is homothetic, the required result holds.

To prove that only a homothetic preference ordering will work, we show that the demand functions exhibit expenditure proportionality. Instead of focusing on the comparison quantities, $H(P^a,\mu^a)$, we focus on the reference quantities, $H(P^b,\mu^b)$. This is simpler because μ^a appears twice in the quantity index while μ^b appears only once. Let P^b be any price vector, and set P^a equal to P^b. Choose μ^a arbitrarily to fix the base indifference curve of the quantity index. Since $P^a = P^b$, the cost-of-living index is 1, and hence the quantity index is equal to μ^a/μ^b for all μ^a,μ^b. Hence, it is homogeneous of degree -1 in μ^b. A slight modification of the argument used in our first proof establishes that $X^b = \lambda X^a$, so the demand functions exhibit expenditure proportionality.

9. Price Indexes for Demand Analysis

Demand theory tells us that the demand for a good is a function of its own price, the prices of all other goods, and total expenditure. Without additional assumptions, the theory says very little about the form of the demand functions. To estimate demand functions, specific assumptions must be made about their functional form. It is sometimes assumed that the demand for each good is a function of its own price and expenditure where these variables have been deflated by an appropriate price index. That is

$$h^i(P,\mu) = g^i \left[\frac{p_i}{T(P)}, \frac{\mu}{T(P)} \right] \qquad (9.1)$$

where the price index T(P) is assumed to be homogeneous of degree 1. The same index is used to deflate both price and expenditure and appears in every demand equation. This means that p_i appears twice in the demand function for the ith good: in its own right as a price variable and again as an argument of the index function T. If the demand func-

tions are of this form, the prices of "other goods" enter the demand functions only through the index function.

In Pollak [1972] I defined "generalized additive separability" as follows: a system of demand functions exhibits generalized additive separability if its demand functions are of the form

$$h^i(Y) = \Gamma^i[y_i, R(Y)], \qquad i = 1,\ldots,n \;.$$

That is, the demand for each good is a function of its own normalized price and an index function which depends on all normalized prices. The same index function appears in the demand function for every good.

Theorem: A system of demand functions is of the form (9.1) if and only if it exhibits generalized additive separability and the index function R(Y) is homothetic.

Proof: If the demand functions are (9.1), we can define $\bar{g}^i(.,.)$ by

$$\bar{g}^1\left[\frac{p_i}{\mu}, \frac{T(P)}{\mu}\right] = g^i\left[\frac{p_i}{T(P)}, \frac{\mu}{T(P)}\right]$$

since the original arguments of g^i can be recovered from those of $\bar{g}^i$. Since the demand functions are homogeneous of degree 0 in P and μ,

$$\bar{g}^i\left[\frac{p_i}{\mu}, \frac{T(P)}{\mu}\right] = \bar{g}^i\left[\frac{\lambda p_i}{\lambda\mu}, \frac{T(\lambda P)}{\lambda\mu}\right] = \bar{g}^i\left[\frac{p_i}{\mu}, \frac{T(\lambda P)/\lambda}{\mu}\right] = \bar{g}^i[y_i, T(Y)] \;,$$

so the demand functions exhibit generalized additive separability where T(Y) is homogeneous of degree 1.

If $h^i(Y) = \Gamma^i[y_i, R(Y)]$, where R is homothetic, we can redefine R and Γ so R is homogeneous of degree 1. Then

$$h^i(P,\mu) = \Gamma^i\left[\frac{p_i}{\mu}, \frac{R(P)}{\mu}\right] = g^i\left[\frac{p_i}{R(P)}, \frac{\mu}{R(P)}\right]$$

since the original arguments of Γ^i can be recovered from those of g^i.

Fourgeaud and Nataf [1959] explicitly characterize the systems of demand functions of the form (9.1) which satisfy the budget constraint and the Slutsky symmetry conditions and hence are theoretically plausible in the sense that they can be derived from a well-behaved preference ordering. Their results are summarized in Pollak [1972, Section IId]. There are four principal cases:

$$h^i(Y) = \frac{\Psi^i[\log y_i - \log R(Y)]}{y_i} \tag{9.2}$$

where $\Psi^i(.)$ is a function of a single variable, and R is defined implicitly by

$$\Sigma\, \Psi^k[\log y_k - \log R] = 1.$$

$$h^i(Y) = \frac{a_i R(Y) + b_i[1 - R(Y)]}{y_i} \tag{9.3}$$

where R(Y) is homogeneous of degree 1 in Y and

$$R(Y) = T[\Sigma\, a_k \log y_k, \Sigma\, b_k \log y_k], \qquad \Sigma\, a_k = \Sigma\, b_k = 1.$$ [14]

The other two cases are

$$h^i(Y) = \frac{\delta[T(Y)][\alpha_i \log y_i - \alpha_i \log T(Y) + \beta_i] + \alpha_i}{y_i}, \tag{9.4}$$

$$\Sigma\, \alpha_k = 1, \Sigma\, \beta_k = 0$$

where $\delta(.)$ is a function of a single variable and

$$T(Y) = \pi\, y_k^{\alpha_k}$$

and

$$h^i(Y) = \frac{\delta[T(Y)][\beta_i - \alpha_i y_i^c T(Y)^{-c}] + \beta_i}{y_i}, \quad \Sigma \beta_k = 1 \tag{9.5}$$

where $\delta(.)$ is a function of a single variable and

$$T(Y) = [\Sigma \alpha_k y_k^c]^{\frac{1}{c}} .$$

The first case, (9.2), exhibits expenditure proportionality and hence is not very interesting for empirical demand analysis. In (9.3) the demand functions are locally linear in expenditure; the index function is a function of two Cobb-Douglas functions. In (9.4) and (9.5) the index functions are Cobb-Douglas and C.E.S., respectively.

Thus, the price indexes which appear in the demand functions depend in a specific way on the parameters of the preference ordering. There is no presumption that these price indexes coincide with the cost-of-living index; indeed, it is not entirely clear what the assertion that $T(P)$ coincides with the cost-of-living index $I(P^a,P^b,s,R)$ would mean. To indicate that the prices in the demand functions are variables, we rewrite (9.1) as

$$h^i(P^t,\mu^t) = g^i\left[\frac{p_i^t}{T(P^t)}, \frac{\mu^t}{T(P^t)}\right]. \tag{9.1$'$}$$

We may interpret P^t as the price vector of period t although this is not essential. The cost-of-living index $I(P^t,P^b,s,R)$ is defined by

$$I(P^t,P^b,s,R) = \frac{E(P^t,s,R)}{E(P^b,s,R)} .$$

Two difficulties are immediately apparent: the first is that the cost-of-living index depends on the reference prices P^b as well as on P^t while the demand functions depend only on the comparison prices P^t. Since the reference prices (and also the base indifference curve) are held constant, the cost-of-living index is proportional to $E(P^t,s,R)$; the constant factor of proportionality is $1/E(P^b,s,R)$, and it can be absorbed into the parameters of the demand functions. Hence, we interpret the assertion that the cost-of-living index coincides with T(P) to mean that

$$h^i(P^t,\mu^t) = g^i\left[\frac{p_i^t}{E(P^t,s)}, \frac{\mu^t}{E(P^t,s)}\right].$$

The fact that the cost-of-living index depends on the choice of the base indifference curve (unless the demand functions exhibit expenditure proportionality) is a more serious problem. In particular, in the absence of expenditure proportionality the assertion that the cost-of-living index coincides with T(P) must be interpreted to mean that there exists a base indifference curve s* such that T(P) is proportional to E(P,s*,R) where the factor of proportionality can be absorbed into T(P).

The linear expenditure system

$$h^i(P,\mu) = b_i - \frac{a_i}{p_i} \Sigma\, b_k p_k + \frac{a_i}{p_i}\mu$$

provides a good illustration. These demand functions can be written as

$$h^i(P,\mu) = b_i - \frac{a_i}{\left(\dfrac{p_i}{\Sigma\, b_k p_k}\right)} + a_i \frac{\left(\dfrac{\mu}{\Sigma\, b_k p_k}\right)}{\left(\dfrac{p_i}{\Sigma\, b_k p_k}\right)} \tag{9.6}$$

which belongs to the Fourgeaud-Nataf class. In fact, the linear expenditure system is a special case of (9.5) where $\delta(T) = -T$ and $c = 1$. The cost-of-living index corresponding to the linear expenditure system is given by (6.2.21). For $s = 0$ this becomes

$$I(P^t,P^b,0,R) = \frac{\Sigma\, b_k p_k^t}{\Sigma\, b_k p_k^b}$$

which is the appropriate price index for demand analysis in this case.

In general, however, the cost-of-living index does not coincide with T(P). For example, the indirect utility function

$$\Psi(P,\mu) = -\frac{1}{\mu}\pi\, p_k^{a_k} + \Sigma\,(a_k - b_k)\log p_k, \qquad \Sigma\, a_k = \Sigma\, b_k = 1$$

is a special case of (9.4). The demand functions are given by

$$h^i(P,\mu) = \frac{a_i}{p_i} - \frac{(a_i - b_i)\mu^2}{p_i\, \pi\, p_k^{a_k}} = \frac{a_i\,(\frac{\mu}{T})}{(\frac{p_i}{T})} - \frac{(a_i - b_i)\,(\frac{\mu}{T})^2}{(\frac{p_i}{T})}$$

where $T(P) = \pi\, p_k^{a_k}$. The expenditure function corresponding to this indirect utility function is

$$\mu = \frac{-\pi\, p_k^{a_k}}{s - \Sigma\,(a_k - b_k)\log p_k}$$

and there is no value of s (independent of prices) for which T(P) is proportional to E(P,s,R).

This is not too surprising. On reflection, there was no reason to expect T(P) to coincide with the cost-of-living index. Different price indexes are needed for different purposes, and the cost-of-living index should not be expected to play a role in demand analysis. Furthermore, even in those cases in which T(P) and the cost-of-living index coincide, this relationship is of no use in empirical demand analysis. The trouble is that we do not start out knowing the cost-of-living index. The cost-of-living index cannot be calculated until the unknown parameters of the demand system have been estimated, and the fact that the unknown price index T(P) is proportional to the unknown expenditure function E(P,s*,R)

does nothing to simplify the estimation problem.

Of course, if we know the cost-of-living index, the situation is very different. In that case, regardless of whether the demand functions involve price indexes, it is possible to calculate the demand functions directly from the cost-of-living index. The cost-of-living index $I^*(P^t,P^b,m,R)$ contains all the information about preferences, and it is fairly straightforward to retrieve this information and find the implied demand functions. To do this we write

$$E[P^t,\Psi(P^b,m)] = mI^*(P^t,P^b,m,R) .$$

Differentiating with respect to p_i^t yields

$$f^i[P^t,\Psi(P^b,m)] = \frac{\partial E[P^t,\Psi(P^b,m)]}{\partial p_i^t} = m \frac{\partial I^*(P^t,P^b,m,R)}{\partial p_i^t} .$$

The ordinary demand functions $h^i(P^t,\mu^t,R)$ can be calculated by finding the value of m, m^t, for which

$$\mu^t = E[P^t,\Psi(P^b,m^t)] = m^t I^*(P^t,P^b,m^t,R) .$$

Then

$$h^i(P^t,\mu^t) = f^i[P^t,\Psi(P^b,m^t)] .$$

If we begin knowing only the cost-of-living index $I(P^t,P^b,s,R)$, then we cannot retrieve the ordinary demand functions. However, even in terms of the fundamental question which the cost-of-living index is designed to answer, it is not enough to know only $I(P^t,P^b,s,R)$. The difficulty is that unless we also know the indirect utility function $\Psi(P,\mu)$, or have other equivalent information, we have no way to interpret the numerical value of s. We cannot associate it with a particular collection of goods and services or with a particular level of expenditure at a particular set of prices. Hence, if we know $I(P^t,P^b,s,R)$ and have enough information to interpret it meaningfully as a cost-of-living index, we can calculate the im-

plied demand functions.

To summarize: it is useful to consider systems of demand functions which involve price indexes, but the price indexes which are relevant for demand analysis are unlikely to coincide with the cost-of-living index. Even if they coincide, knowing this is of no help in empirical demand analysis because the cost-of-living index cannot be computed until the unknown parameters of the system of demand equations have been estimated. The cost-of-living index formula $I^*(P^t,P^b,m,R)$ contains enough information about preferences to enable us to calculate the demand functions. So, if the formula for the cost-of-living index is known, then the demand functions can be calculated from it, regardless of whether the demand functions involve price indexes.

We now turn to more practical questions about the use of price indexes in empirical demand analysis. Assuming that the demand functions depend on price indexes, (9.1), when will the index T(P) be a linear function of prices? We have already seen that the linear expenditure system is of this form; we now characterize the entire class.

Theorem: If the demand functions are of the Fourgeaud-Nataf form (9.1)

$$h^i(P,\mu) = g^i\left[\frac{p_i}{T(P)}, \frac{\mu}{T(P)}\right], \qquad T(\lambda P) = \lambda T(P)$$

and if the index function T(P) is linear in prices

$$T(P) = \Sigma\, w_k p_k\,,$$

then, except for degenerate cases in which less than three w's are non-zero, the demand functions fall into two classes:

$$h^i(P,\mu) = \frac{\beta_i \mu}{p_i} + \frac{(1 - \Sigma\, \beta_k)\alpha_i \mu}{\Sigma\, \alpha_k p_k} \qquad (9.7)$$

and

$$h^i(P,\mu) = \frac{\beta_i \mu}{p_i} + \delta\left[\frac{\Sigma\, \alpha_k p_k}{\mu}\right]\left[\frac{\beta_i \mu}{p_i} - \frac{\alpha_i}{\Sigma\, \alpha_k p_k}\right], \tag{9.8}$$

where $\delta(.)$ is a function of a single variable.

Proof: We sketch the proof in four parts. (1) First, we show that (9.7) is the only non-degenerate admissible case of (9.2). Differentiating

$$\Sigma\, \Psi^k[\log y_k - \log \Sigma\, w_k y_k] = 1$$

with respect to y_i, we find

$$\frac{\Psi^{i\prime}}{y_i w_i} = \frac{\Psi^{j\prime}}{y_j w_j}.$$

Differentiating with respect to y_s yields

$$\frac{\Psi^{i\prime\prime}}{\Psi^{i\prime}} = \frac{\Psi^{j\prime\prime}}{\Psi^{j\prime}} = c$$

where c is a constant. It is easy to show that $c = 0$ corresponds to a degenerate case. If $c \neq 0$, then

$$\Psi^i(z_i) = \beta_i + \frac{\alpha_i}{c}$$

and the demand functions are given by (9.7). (2) There are no non-degenerate cases corresponding to (9.3). To show this we differentiate

$$\Sigma\, w_k y_k = T[\Sigma\, a_k \log y_k, \Sigma\, b_k \log y_k], \qquad \Sigma\, a_k = \Sigma\, b_k = 1$$

with respect to y_i

$$c_i y_i = a_i T_1 + b_i T_2 \qquad i = 1,\ldots,n$$

and observe that

$$\Sigma\, c_k y_k = T_1 + T_2 = T\,.$$

If two of the equations

$$c_i y_i = a_i T_1 + b_i T_2$$

are independent, we can solve for T_1 and T_2 in terms of the two corresponding y's. Hence, T depends only on those two y's, so this is a degenerate case. If all of the equations are linearly dependent, and some $c_i \neq 0$, we can take $c_1 \neq 0$ with no loss of generality. Then

$$a_i = \lambda a, \qquad b_i = \lambda b, \qquad c_i y_i = \lambda c_1 y_1\,.$$

The last equation implies $\lambda = 0$, and hence $c_i = 0$, $i = 2,\ldots,n$. So $T(Y) = c_1 y_1$. (3) The only admissible case of (9.4) is a degenerate one in which only one is non-zero. (4) The non-degenerate admissible cases of (9.5) are those in which $c = 1$, which is (9.8).

We now consider two other price indexes which might be used to deflate p_i and μ in (9.1′): The Laspeyres and Paasche. Deflating by the Laspeyres index implies

$$T(P^t) = \frac{\Sigma\, h^k(P^b,\mu^b)p_k^t}{\mu^b}.$$

Since P^b and μ^b are constants, this is equivalent to requiring $T(P^t)$ to be a linear function of prices where the weights are proportional to reference period consumption.

This could happen in two ways. First, it might be that the demand functions are such that in every price expenditure situation the quantities demanded are proportional to the α's. That is,

$$\alpha_i = \lambda(P,\mu)h^i(P,\mu) \qquad i = 1,\ldots,n \, .$$

Multiplying by p_i and summing over all goods, we find

$$\lambda(P,\mu) = \frac{\mu}{\Sigma\, \alpha_k p_k}$$

so the demand functions are of the homogeneous fixed coefficient form

$$h^i(P,\mu) = \frac{\alpha_i \mu}{\Sigma\, \alpha_k p_k} \, .$$

This proves that it is only in the homogeneous fixed coefficient case that the reference period consumption pattern is proportional to the α's for all price-expenditure situations. Second, even if every price-expenditure situation does not imply quantities demanded which are proportional to the α's , there are likely to be some price-expenditure situations in which proportionality holds. If, by a fortunate coincidence, the reference price-expenditure situation were one which implied proportionality, then deflation of price and expenditure by the Laspeyres index would be appropriate. There is little to say about the likelihood of this coincidence occurring, but even if it did occur, it is not clear how one would recognize it.

Deflating by the Paasche index is equivalent to setting

$$T(P^t) = \frac{\lambda(P^b)\mu^t}{\Sigma\, h^k(P^t,\mu^t)p_k^b}$$

since we require only proportionality. The factor of proportionality $\lambda(P^b)$ must be independent of both P^t and μ^t. As in the Laspeyres case, this may hold as an identity for all possible reference prices, or if it holds only for some reference prices, it may by coincidence hold for the particular situation we have observed. If it holds for all reference prices, then

$$T(P^t)\, \Sigma\, h^k(P^t,\mu^t)p_k^b = \lambda(P^b)\mu^t$$

is an identity in P^b. Differentiating with respect to p_i^b yields

$$T(P^t)\, h^i(P^t,\mu^t) = \lambda_i(P^b)\mu^t \; .$$

This implies that $\lambda_i(P^b)$ is a constant, say α_i, and so

$$h^i(P^t,\mu^t) = \frac{\alpha_i \mu^t}{T(P^t)} \; .$$

Multiplying by p_i and summing over all goods, we find

$$T(P^t) = \Sigma\, \alpha_k p_k^t$$

which implies that the demand functions are those of the homogeneous fixed coefficient case.

We summarize these results formally:

Theorem: If the demand functions are of the Fourgeaud-Nataf form (9.1)

$$h^i(P,\mu) = g^i \left[\frac{p_i}{T(P)}, \frac{\mu}{T(P)}\right], \qquad T(\lambda P) = \lambda T(P)$$

and if the index function T(P) is proportional to either the Laspeyres or the Paasche index for all reference prices, then the direct utility function is of the homogeneous fixed coefficient form

$$U(X) = \min \left\{\frac{x_i}{a_i}\right\}$$

and the demand functions are of the form

$$h^i(P,\mu) = \frac{a_i \mu}{\Sigma\, a_k p_k}.$$

This means that neither the Laspeyres nor the Paasche index has any special status in demand analysis. In particular, there is no presumption that deflating price and expenditure by the Laspeyres index is better than deflating by a weighted average of prices whose weights were chosen with the aid of a table of random digits. By the same token, there is no presumption that deflation of U.S. food prices and expenditure by a Laspeyres index based on U.S. consumption patterns is better, from a standpoint of demand analysis, than deflating by one whose weights reflect the consumption pattern of Outer Mongolia. The only way to find the appropriate weights is to estimate them along with the other unknown parameters.

Footnotes

[1] (added, 1982) For excellent recent surveys of the literature, see Diewert [1981, 1983]. A number of topics not treated here are discussed in my subsequent papers on cost-of-living indexes: Pollak, [1975a, 1975b, 1978, 1980, 1981, 1983], and Pollak and Wales [1979].

[2] (added, 1982) In the original version, μ was called ''income''; somewhat inconsistently, I have retained such traditional phrases as ''income elasticity'' and ''income-consumption'' curve.

[3] (added, 1982) It is often, although not always, more convenient to denote a particular indifference curve by specifying a commodity vector X which lies on the indifference curve.

[4] (added, 1982) In the original 1971 version of this paper the expenditure function E(P,s) was called the ''cost function'' and denoted by C(P,s). The new terminology is especially convenient in the household production context where the term ''cost function'' is used in a different sense. See Pollak [1978]. It is often convenient to write the expenditure function as E(P,X), where X identifies the base indifference curve.

[5] (added, 1982). I now believe long-run pseudo preferences are inappropriate for welfare comparisons. See Pollak [1976].

[6] (added, 1982). In fact, these results do not depend on preferences satisfying the usual regularity conditions.

[7] (added, 1982). The procedure she describes in the published version of her paper, McElroy [1975] is more general than this sentence suggests.

[8] Some of the material in 6.1 and 6.2 is taken from Pollak [1971].

[9] The requirement $\Sigma a_k = 1$ is a normalization rule and involves no loss of generality.

[10] (added, 1982). This ''generalized Cobb-Douglas'' is a different form from Diewert's [1974, p.116].

[11] However, as Afriat [1972] points out, there are problems with regularity conditions in the homogeneous quadratic case.

[12] ''Real wages'' are sometimes used as a measure of quantity, but they are also used in empirical analysis of the supply or demand for labor. The appropriateness of using an index of real wages obtained in this way for empirical analysis is not explicitly discussed in this paper although the use of price indexes to deflate prices and expenditure in demand analysis is discussed in Section 9. The conclusion there is that such deflation is inappropriate, and it seems clear that deflation of money wages is no better.

[13] In Section 6 we used $\phi(P)$ to denote a function homogeneous of degree 1; in this section ϕ is not assumed to be homogeneous.

[14] There are further restrictions of the function T which we can safely ignore.

* This research was supported in part by the Bureau of Labor Statistics and the National Science Foundation. I am grateful to Irving B. Kravis and Karl Shell for helpful discussions, but neither they, the University of Pennsylvania, the Bureau of Labor Statistics, nor the National Science Foundation should be held responsible for the views expressed.

References

Afriat, S.N. "Theory of International Comparison of Real Income and Prices", In D.J. Daly, ed., *International Comparisons of Prices and Output*. New York: National Bureau of Economic Research, Columbia University Press, 1972.

Fisher, Irving. *The Making of Index Numbers*. Cambridge, Massachusetts: Houghton Mifflin Co., 1922.

Fisher, F.M. and K. Shell. "Taste and Quality Change in the Pure Theory of the True Cost-of-Living Index". In J.N. Wolfe, ed., *Value, Capital and Growth: Papers in Honour of Sir John Hicks*. Edinburgh: University of Edinburgh Press, 1968.

Fourgeaud, C. and A. Nataf. "Consommation en Prix et Revenu Reels et Theorie des Choix", *Econometrica*, Vol. 27, No. 3 (July 1959), pp. 329-354.

Gorman, W.M. "On a Class of Preference Fields", *Metroeconomica*, Vol. 13 (August 1961), pp. 53-56.

Klein, L.R. and H. Rubin. "A Constant-Utility Index of the Cost-of-Living", *Review of Economic Studies*, Vol. XV (2), No. 38 (1947-48), pp. 84-87.

Malmquist, S. "Index Numbers and Indifference Surfaces", *Trabajos de Estadistica* (1953), 4, pp. 209-242.

McElroy, Marjorie. "A 'Kinked' Linear Expenditure System: An Application to Rural South Vietnamese Households", mimeograph, 1969.

Pollak, Robert A. "Additive Utility Functions and Linear Engel Curves", *Review of Economic Studies*, Vol. 38, No. 4 (October 1971), pp. 401-414.

__________ "Generalized Separability", *Econometrica*, Vol. 40, No. 3 (May 1972), pp. 431-453.

_________ "Habit Formation and Dynamic Demand Functions", *Journal of Political Economy*, Vol. 78, No. 4 (July-August 1970), pp. 745-763.

Samuelson, Paul A. *Foundations of Economic Analysis*. Cambridge, Massachusetts: Harvard University Press, 1947.

Wegge, L.L. "The Demand Curves from a Quadratic Utility Indicator", *Review of Economic Studies*, Vol. XXXV (2), No. 102 (April 1968), pp. 209-224.

Wold, H.O.A. and L. Jureen. *Demand Analysis*. New York: John Wiley and Sons, 1953.

Additional References, 1982

Diewert, W.E., ''Applications of Duality Theory'', in M. Intriligator and D. Kendrick, eds., *Frontiers of Quantitative Economics II*, Amsterdam: North-Holland Publishing Co., 1974, pp. 106-171.

_________ ''The Economic Theory of Index Numbers: A Survey'', in A. Deaton, ed., *Essays in the Theory and Measurement of Consumer Behaviour in Honour of Sir Richard Stone*. London: Cambridge University Press, 1981, pp. 163-208.

_________ ''The Theory of the Cost-of-Living Index and the Measurement of Welfare Change'', *Price Level Measurement: Proceedings From a Conference Sponsored by Statistics Canada,* (1983).

McElroy, Marjorie, ''A Spliced CES Expenditure System'', *International Economic Review*, Vol. 16, No. 3 (October 1975), pp. 765-780.

Pollak, Robert A., ''Subindexes of the Cost-of-Living Index'', *International Economic Review*, Vol. 16, No. 1 (February 1975), pp. 135-150. (1975a)

_________ ''The Intertemporal Cost-of-Living Index'', *Annals of Economic and Social Measurement*, Vol. 4, No. 1, (Winter 1975), pp. 179-195. (1975b)

_________ ''Habit Formation and Long-Run Utility Functions'', *Journal of Economic Theory*, Vol. 13, No. 2 (October 1976), pp. 272-297.

_________ ''Welfare Evaluation and the Cost-of-Living Index in the Household Production Model'', *American Economic Review*, Vol. 68, No. 3, (June 1978), pp. 285-299.

_________ ''Group Cost-of-Living Indexes'', *American Economic Review*, Vol. 70, No. 2 (May 1980), pp. 273-278.

_________ ''The Social Cost-of-Living Index'', *Journal of Public Economics*, Vol. 15, No. 3, (June 1981), pp. 311-336.

_________ ''The Treatment of 'Quality' in the Cost-of-Living Index'', *Journal of Public Economics*, Vol. 20, No. 1, (February 1983), pp.25-53.

Pollak, Robert A., and Terence J. Wales, ''Welfare Comparisons and Equivalence Scales'', *American Economic Review*, Vol. 69, No. 2, (May 1979), pp. 216-221.

THE THEORY OF THE COST-OF-LIVING INDEX AND THE MEASUREMENT OF WELFARE CHANGE

*W.E. Diewert**
Department of Economics
University of British Columbia

SUMMARY

The Consumer Price Index is often regarded as an approximation to a Cost-of-Living Index. This paper reviews the theoretical foundations of the Cost-of-Living Index and the closely related problems involved in measuring changes in economic welfare.

The Cost-of-Living Index for a single person is defined as the minimum cost of achieving a certain standard of living during a given period divided by the minimum cost of achieving the same standard of living during a base period. In order to numerically construct an individual's Cost-of-Living Index, it is necessary to know his or her preferences over economic goods. Since these preferences are essentially unobservable, it is necessary to construct approximations to the Cost-of-Living Index. This topic is discussed in Section 2 of the paper.

The remaining sections of the paper discuss a number of related topics, including: the closely related problems involved in measuring a group Cost-of-living index and changes in the welfare of a group, the fixed based versus the chain principle, the choice of a functional form for the Cost-of-Living Index, the treatment of durable goods, such as housing and the treatment of taxes and labour supply in a Cost-of-Living Index.

RÉSUMÉ

L'indice des prix à la consommation est souvent perçu comme une approximation de l'indice du coût de la vie. Cette communication passe en revue les fondements théoriques

de l'indice du coût de la vie et les problèmes étroitement connexes liés à la mesure des variations du bien-être économique.

L'indice du coût de la vie pour une personne seule se définit comme le coût minimum associé à un certain niveau de vie au cours d'une période donnée divisé par le coût minimum du même niveau de vie au cours d'une période de base. Afin de construire numériquement l'indice du coût de la vie d'un particulier, il faut connaître ses préférences quant aux biens économiques. Puisque ces préférences sont essentiellement inobservables, il y a lieu de construire des approximations de l'indice du coût de la vie. C'est cette question qui fait l'objet de la section 2 de la présente communication.

À la section 3, l'auteur étudie le problème très connexe d'obtenir des indicateurs de la variation d'utilité ou de bien-être pour un ménage individuel est étudié.

La section 4 traite du principe de l'enchaînement utilisé dans la construction des nombres-indices comme l'indice des prix à la consommation et le met en contraste avec le principe de la base fixe.

Les sections 5 et 6 traitent d'autres concepts possibles pour un indice collectif du coût de la vie. **L'indice de prix démocratique** *est une moyenne simple des indices du coût de la vie des ménages individuels.* **L'indice du coût social de la vie** *est défini par Prais et Pollak comme la dépense minimale requise pour atteindre un certain niveau de bien-être pour chaque ménage de l'économie lorsque l'économie fait face à des prix de période 1 par rapport à la dépense minimale requise pour atteindre le même niveau de bien-être pour chaque ménage lorsque l'économie fait face à des prix de période 0. Afin d'établir des approximations de l'indice de prix démocratique, il faut disposer de données sur les consommateurs individuels. Cependant, on peut obtenir des approximations du coût social de la vie en utilisant des données agrégatives.*

La section 7 traite des problèmes épineux que comporte la formation d'indices collectifs de bien-être.

La section 8 aborde les questions suivantes: (i) comment peut-on définir un sous-indice

de l'indice du coût de la vie, et (ii) comment faut-il regrouper les sous-indices pour obtenir une approximation de l'indice global du coût de la vie?

Comment faut-il traiter les économies et les questions de finances des consommateurs dans l'indice du coût de la vie? Ces questions font surface à la section 9, qui traite de la théorie des indices intertemporels du coût de la vie. Cependant, même si l'on peut étudier théoriquement les indices intertemporels, ils sont difficiles à réaliser en pratique, puisqu'ils dépendent des attentes des ménages à l'égard des prix à venir, qui sont essentiellement inobservables.

Les indices spatiaux du coût de la vie, qui comparent le niveau des prix d'un endroit géographique avec celui d'un autre endroit au même moment, font une brève apparition à la section 10.

La section 11 présente un examen plus approfondi du problème de la modélisation de l'offre de loisirs et de main-d'oeuvre dans un indice du coût de la vie.

La section 12 traite du problème de la modélisation des biens durables (comme le logement et l'automobile) dans l'indice du coût de la vie. Elle développe l'approche du coût d'utilisation à l'égard de ce problème. Les attentes du ménage pour ce qui est des prix à venir, la situation financière du ménage (est-il emprunteur ou prêteur, et à quels taux), et la situation du ménage face à l'impôt sur le revenu sont autant d'éléments qui jouent un rôle crucial dans cette approche du coût d'utilisation. Ainsi, les données qu'exige cette approche sont malheureusement très volumineuses.

La section 13 présente quelques brèves observations sur le problème des nouveaux biens.

La section 14 conclut avec quelques recommandations.

La section 15 est une annexe qui présente les preuves des nouveaux résultats théoriques.

1. Introduction

As the title of the paper indicates, we will investigate the theoretical foundations of the

cost-of-living index. This seems appropriate in a conference about the Consumer Price Index (CPI), since the CPI is now being used as a proxy for the cost-of-living index in indexing contracts and as an inflation measure.[1] We shall also discuss the closely related issues involved in measuring welfare changes, both for individual consumers and for groups of consumers.

The economic theory of the cost-of-living index for a single household is reviewed in Section 2.

In Section 3, we study the closely related problem of obtaining single household indicators of utility or real income change.

In Section 4, we discuss the costs and benefits of using the chain principle in the construction of index number formulae versus the fixed base principle.

In Sections 5 to 7, we discuss various concepts that have been proposed for group cost-of-living and welfare indexes.

In Sections 8 and 9, we outline the theory of subindexes of the cost-of-living index and the related idea of an intertemporal cost-of-living index.

In Section 10, we consider the problem of constructing price indexes that compare the level of prices in different locations.

Labour and durable goods in the cost-of-living index make their appearance in Sections 11 and 12 respectively.

Section 13 discusses the new goods problem and Section 14 concludes with some recommendations.

In a companion paper, Diewert [1983], we discuss price and output indexes from the viewpoint of producer theory. The reader may be aware of the old Hicks [1940; 1958; 1981] - Samuelson [1950; 1961] measurement of real income controversy; i.e., is there such an

animal as real income, and if so, should it be measured from the consumer or producer point of view? From the consumer point of view, our conclusion is that real income is a very subjective animal and hence it probably does not exist, unless we are willing to give explicit numerical weights to the welfares of different household classes. On the producer side, the situation is more encouraging: although real output is not a useful concept, the closely related concept of total factor productivity does turn out to be useful. For the details of this "new" approach to measurement total factor productivity (which in fact was suggested many years ago by Hicks [1961] [1981; 192-3]), see Caves, Christensen and Diewert [1982b] and Diewert [1983].

Section 15 is an Appendix that collects proofs of new theorems.[2]

2. The Single Household Cost-of-Living Index

We assume that the household or individual has recurring preferences over combinations of N goods that may be represented by a utility function F where $u = F(x)$ is the utility level or standard of living that can be attained if the individual consumes the consumption vector $x \equiv (x_1, x_2, \ldots, x_N)^T \geq 0_N$.[3]

We assume that the utility function F satisfies Conditions I which are technical enough to relegate to a footnote.[4]

We shall assume that the consumer maximizes his utility function F(x) subject to a budget constraint of the form $p \cdot x = \Sigma_{n=1}^{N} p_n x_n \leq y$ where $p >> 0_N$ is a positive vector of commodity (rental) prices and $y > 0$ is expenditure on the N commodities.

The consumer's utility maximization problem can be decomposed into two stages. In the first stage, the consumer attempts to minimize the cost of achieving a given utility level, and in the second stage, he chooses the maximal utility level that is just consistent with his budget constraint.

The solution to the first-stage problem defines the consumer's *cost function* C: for $u \geq 0$, $p >> 0_N$

$$C(u,p) \equiv \min_x \{p \cdot x : F(x) \geq u, x \geq 0_N\} \quad . \tag{1}$$

Given that F satisfies Conditions I, C will satisfy Conditions II (which we relegate to another footnote[5]). Moreover, if we are given a cost function C satisfying Conditions II, C may be used in order to construct the underlying preference function F which will satisfy Conditions I.[6]

Our interest in C stems from the fact that it may be used to define the Konüs [1924] *cost-of-living index* P_K: for $p^0 >> 0_N$, $p^1 >> 0_N$ and $u > 0$ define

$$P_K(p^0, p^1, u) \equiv C(u, p^1)/C(u, p^0) \quad . \tag{2}$$

Thus P_K depends on three variables: (i) p^0, a vector of period 0 or base period prices, (ii) p^1, a vector of period 1 or current period prices, and (iii) u, a number that indexes the reference indifference surface. Thus $P_K(p^0, p^1, u)$ is the minimum cost of achieving the standard of living indexed by u when the consumer faces period 1 prices p^1 **relative** to the minimum cost of achieving the same standard of living when the consumer faces period 0 prices p^0. If there is only one good, then it can be seen that $P_K(p_1^0, p_1^1, u) = p_1^1/p_1^0$ for all $u > 0$. In this case, there is obviously no index number problem.

In the general case when there is more than one good, the functional form for the cost-of-living index P_K obviously depends on the functional form for the consumer's cost function C, which in turn is determined by the form of the consumer's preference function F. Our fundamental problem is that we do not know what the functional forms for F or C and hence P_K are. Our primary task in this section will be to see if we can find adequate bounds or approximations to the true cost-of-living index P_K that depend only on observable market price and quantity data. However, before we turn to this primary task, we state a theorem which provides necessary and sufficient conditions for a given function $P(p^0, p^1, u)$ of $2N+1$ variables to be interpretable as a cost-of-living index.

Theorem 1

Let P be a function of 2N + 1 positive variables that satisfies the following properties: for all $u > 0$, $p^0 >> 0_N$, $p^1 >> 0_N$, and $p^2 >> 0_N$, we have (i) $P(p^0, p^1, u) > 0$ (positivity), (ii) $P(p^0, p^1, u) = 1/P(p^1, p^0, u)$ (time reversal property), (iii) $P(p^0, p^2, u) = P(p^0, p^1, u) P(p^1, p^2, u)$ (circularity or transitivity) and (iv) for some $p^* >> 0_N$, $C(u,p) \equiv uP(p^*, p, u)$ regarded as a function of u and p satisfies Conditions II for a cost function. Then P is the cost-of-living index that corresponds to the preferences that are dual to C; i.e., $P \equiv P_K$ satisfies(2). Moreover, C satisfies the following money metric[7] scaling of utility property:

$$C(u, p^*) = u \text{ for all } u > 0. \qquad (3)$$

Conversely, given a cost function C satisfying Conditions II and the money metric property (3), then $P \equiv P_K$ defined by (2) satisfies properties (i) to (iv) listed above in the theorem.

The above theorem is very closely related to some results in Pollak [1983], who stressed that the mathematical properties of P_K are completely characterized by the mathematical properties of C.

The following theorem[8] provides observable bounds on $P_K(p^0, p^1, u)$.

Theorem 2

(Lerner [1935-36], Joseph [1935-36], Samuelson [1947; p.159]). P_K lies between the lowest price ratio and the highest price ratio for any reference indifference surface indexed by $u > 0$; i.e.,

$$\begin{aligned} \min_i \{p_i^1/p_i^0 : i = 1, \ldots, N\} &\leq P_K(p^0, p^1, u) \\ &\leq \max_i \{p_i^1/p_i^0 : i = 1, \ldots, N\}. \end{aligned} \qquad (4)$$

The above limits are wide, but they are not useless. For example, if prices vary in strict proportion so that $p^1 = \lambda p^0$ for some $\lambda > 0$, then the upper and lower limit in (4) is λ and hence $P_K(p^0, \lambda p^0, u) = \lambda$ also.

In order to make further progress, we shall assume cost minimizing behaviour on the part of the consumer during periods 0 and 1. We shall also assume that we can observe the consumer's quantity choices $x^0 > 0_N$ and $x^1 > 0_N$ made during periods 0 and 1 in addition to the corresponding price vectors $p^0 >> 0_N$ and $p^1 >> 0_N$. Thus we assume:

$$p^0 \cdot x^0 = C[F(x^0), p^0] \; ; \; p^1 \cdot x^1 = C[F(x^1), p^1]. \tag{5}$$

We have introduced the concept of the Konüs cost-of-living index $P_K(p^0, p^1, u)$ without saying much about the choice of the reference indifference surface indexed by u. It would appear that there are two natural choices for u: namely $F(x^0)$ or $F(x^1)$. Thus the *Laspeyres-Konüs cost-of-living index* is defined as:

$$P_K(p^0, p^1, F(x^0)) \equiv C[F(x^0), p^1]/C[F(x^0), p^0] \tag{6}$$

while the *Paasche-Konüs cost-of-living index* is defined as:

$$P_K(p^0, p^1, F(x^1)) \equiv C[F(x^1), p^1]/C[F(x^1), p^0]. \tag{7}$$

In order to understand why the indexes (6) and (7) are named after Laspeyres and Paasche it is first necessary to introduce the concept of a *mechanical price index formula*. This is simply a function P of the observable price and quantity vectors for the two periods, p^0, p^1, x^0, x^1, of known functional form. Two examples of such formulae are the *Laspeyres price index* P_L defined by

$$P_L(p^0, p^1, x^0, x^1) \equiv p^1 \cdot x^0 / p^0 \cdot x^0 \tag{8}$$

and the *Paasche price index* P_P defined by

$$P_P(p^0,p^1,x^0,x^1) \equiv p^1 \cdot x^1 / p^0 \cdot x^1. \tag{9}$$

Irving Fisher [1922] gives hundreds of examples of mechanical price index formulae. The axiomatic characterization of these indexes may be found in Eichhorn [1976; 1978] and Eichhorn and Voeller [1976; 1983]. Note that these indexes are functions of 4N arguments whereas the price index that appeared in Theorem 1 had only 2N+1 arguments.

The following theorem relates the (unobservable) Laspeyres-Konüs cost-of-living index defined by (6) to the (observable) Laspeyres price index defined by (8), and the Paasche-Konüs cost-of-living index defined by (7) to the Paasche price index defined by (9).

Theorem 3

(Konüs [1924; pp.17-19]): Assuming (5), cost minimizing behaviour during periods 0 and 1, we have:

$$P_K(p^0,p^1,F(x^0)) \leq p^1 \cdot x^0 / p^0 \cdot x^0 \equiv P_L; \tag{10}$$

$$P_K(p^0,p^1, F(x^1)) \geq p^1 \cdot x^1 / p^0 \cdot x^1 \equiv P_P \, . \tag{11}$$

Corollary (Pollak [1983]):

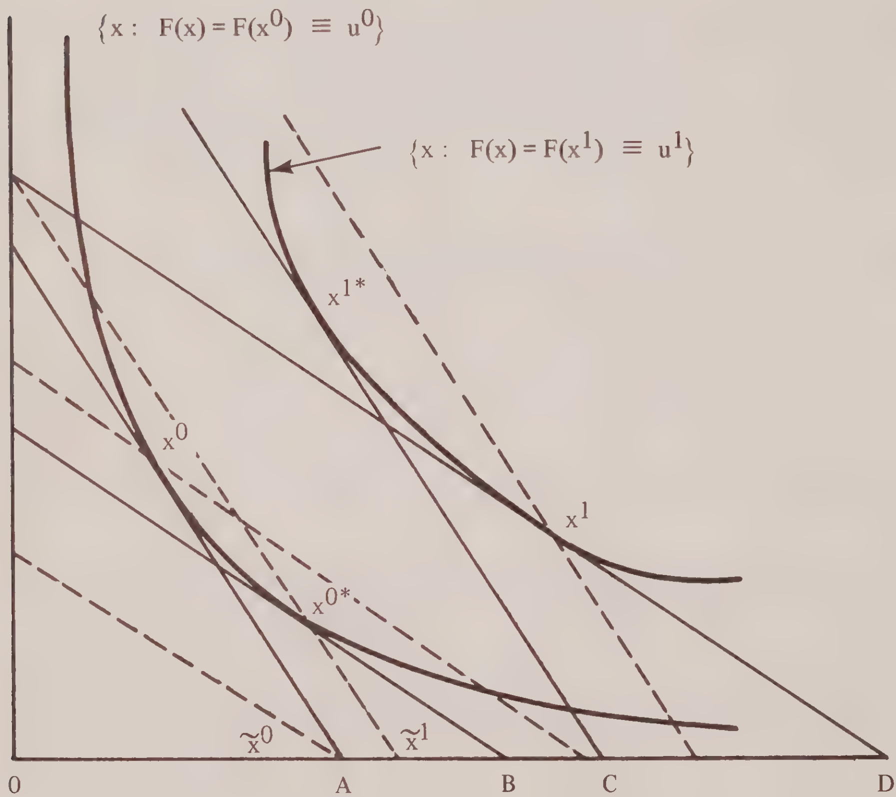
$$\min_i\{p_i^1/p_i^0\} \leq P_K(p^0,p^1,F(x^0)) \leq P_L \equiv p^1 \cdot x^0 / p^0 \cdot x^0; \tag{12}$$

$$P_P \equiv p^1 \cdot x^1 / p^0 \cdot x^1 \leq P_K(p^0,p^1, F(x^1)) \leq \max_i\{p_i^1/p_i^0\}. \tag{13}$$

The above corollary follows combining Theorems 2 and 3. The Laspeyres-Konüs index $P_K(p^0,p^1,F(x^0)) \equiv C[F(x^0),p^1]/C[F(x^0),p^0] = p^1 \cdot x^{0*} / p^0 \cdot x^0$ is illustrated in Figure 1 in the two good case along with the bounds in (12). Note that x^{0*} is the solution to the problem of minimizing the cost of achieving the utility level $u^0 \equiv F(x^0)$ when the consumer is faced with period 1 prices p^1. Although $p^1 \cdot x^{0*}$ is not observable, the upper bound $p^1 \cdot x^0$

(see the dashed straight line through x^0) is observable. The lower bound to $p^1 \cdot x^{0*}$ is $p^1 \cdot \tilde{x}^0$, where $\tilde{x}^0$ is the point where the budget line $\{x: p^0 \cdot x = p^0 \cdot x^0\}$ intersects the x_1 axis, and it too is observable (see the dashed line through $\tilde{x}^0$). These upper and lower bounds correspond to the upper and lower bounds in (12). Analytically, the upper bound rests on the fact that the point x^0 is an **inner approximation** to the consumer's true indifference set $L(u^0) \equiv \{x: x \geq F(x^0) \equiv u^0\}$ while the set that is on or above the period 0 budget line $x: p^0 \cdot x \geq p^0 \cdot x^0$ intersected with the non-negative orthant x: $\{x: x \geq 0_N\}$ is an **outer approximation** to the true indifference set $L(u^0)$.

Figure 1.



A similar analysis can be carried out for the Paasche-Konüs price index $P_K(p^0,p^1,F(x^1)) \equiv C[F(x^1), p^1]/C[F(x^1),p^0] = p^1 \cdot x^1/p^0 \cdot x^{1*}$. From Figure 1, it can be seen that $p^0 \cdot x^1$ is an upper bound for $p^0 \cdot x^{1*}$ while $p^0 \cdot \tilde{x}^1$ is a lower bound for $p^0 \cdot x^{1*}$. These upper and lower bounds for $p^0 \cdot x^{1*}$ yield the lower and upper bounds in (13).

It can be seen from Figure 1, that the Laspeyres index P_L will be rather close to $P_K(p^0,p^1,F(x^0))$ provided that the indifference surface x: $F(x)=F(x^0)$ is not too linear around x^0. (The perfectly linear case corresponds to the perfect substitutes case). Similarly, the Paasche index P_P will be close to $P_K(p^0, p^1,F(x^1))$ provided that the indifference surface $x:\{F(x)=F(x^1)\}$ is not too linear around x^1. This is an encouraging observation, but it still does not tell us how close P_L is to $P_K(p^0,p^1,F(x^0))$ or how close P_P is to $P_K(p^0,p^1,F(x^1))$.

In order to make further progress, we may proceed in three directions: (i) introduce additional observations $(p^2,x^2),\ldots,(p^T,x^T)$ and use the revealed preference techniques associated with Samuelson [1947] and Afriat [1967; 1972; 1979] in order to form nonparametric approximations to the consumer's preferences,[9] (ii) make specific functional form assumptions about F or C, or (iii) choose the reference utility level u that occurs in $P_K(p^0,p^1,u)$ in an empirically convenient manner. We will pursue only possibilities (ii) and (iii) in this paper.

It is an empirical fact that the Laspeyres and Paasche indexes, $P_L(p^0,p^1,x^0,x^1)$ and $P_P(p^0,p^1,x^0,x^1)$, are often rather close to each other numerically (we will return to this point in Section 4). Thus the following theorem is extremely important from a practical point of view.

Theorem 4

(Konüs [1924; pp.20-21]): Let the consumer's utility function F satisfy Conditions I and suppose that the observed data for periods 0 and 1, (p^0,x^0) and (p^1,x^1) respectively, satisfy the cost minimization assumptions (5). Then there exists a reference utility level u* that lies between[10] the base utility level $u^0 \equiv F(x^0)$ and the period 1 utility level $u^1 \equiv F(x^1)$ such that the consumer's true cost-of-living index for this reference utility level, $P_K(p^0,p^1,u^*)$, lies between $P_L \equiv p^1 \cdot x^0/p^0 \cdot x^0$ and $P_P \equiv p^1 \cdot x^1/p^0 \cdot x^1$; i.e., we have

$$P_P \leq P_K(p^0,p^1,u^*) \leq P_L \qquad \text{if } P_P \leq P_L \text{ or} \tag{14}$$

$$P_L \le P_K(p^0,p^1,u^*) \le P_P \qquad \text{if } P_L \le P_P . \tag{15}$$

In most applications of index number theory, we would be quite happy if we knew $P_K(p^0,p^1,u^0)$ or $P_K(p^0,p^1,u^1)$ or $P_K(p^0,p^1,u^*)$ for some u^* between u^0 and u^1. Hence if P_L is numerically close to P_P, $P_P(p^0,p^1,u^*)$ will be squeezed in by these two numbers and we will have the consumer's true cost of living between periods 0 and 1 for all practical purposes.

It would be pleasant if we could extend Theorem 4 to conclude that there exists a u^* between u^0 and u^1 such that $P_K(p^0,p^1,u^*)$ equals some specific average of P_L and P_P, such as Irving Fisher's [1922] ideal index number formula P_2 which is defined as the geometric average of P_L and P_P; i.e.,

$$P_2(p^0,p^1,x^0,x^1) \equiv [p^1 \cdot x^0/p^0 \cdot x^0]^{1/2}[p^1 \cdot x^1/p^0 \cdot x^1]^{1/2}.$$

Unfortunately, we cannot draw such a conclusion in general. The problem is that $P_K(p^0,p^1,u)$ could be rather close to **either** P_L **or** P_P for all u between u^0 and u^1 and hence if P_L and P_P are rather different, then their geometric mean P_2 could lie above or below $P_K(p^0,p^1,u)$ for all u between u^0 and u^1. However, if P_L and P_P are "close" to each other, then there exists a u^* between u^0 and u^1 such that $P_K(p^0,p^1,u^*)$ is "close" to the Fisher index P_2. This provides a somewhat informal justification for the use of P_2 as an approximation to P_K.

A more formal justification for the use of P_2 that rests on a specific functional form for C(u,p) is also possible. Let us suppose that $C(u,p) = uc^r(p)$ where c^r is defined by:

$$c^r(p) \equiv [\sum_{i=1}^{N} \sum_{j=1}^{N} b_{ij} p_i^{r/2} p_j^{r/2}]^{1/r} \text{ for } r \neq 0, \tag{16}$$

where the parameters b_{ij} satisfy the restrictions $b_{ij} = b_{ji}$ for $1 \le i < j \le N$. The unit cost function c^r defined by (16) is Denny's [1974] quadratic mean of order r unit cost function

which can provide a second order approximation to an arbitrary twice continuously differentiable unit cost function.[11]

Define the base period expenditure shares $s_i^0 \equiv p_i^0 x_i^0 / p^0 \cdot x^0$ and the period 1 expenditure shares $s_i^1 \equiv p_i^1 x_i^1 / p^1 \cdot x^1$ for $i=1,\ldots,N$. For each $r \neq 0$, define the mechanical price index formula P_r by:

$$P_r(p^0,p^1,x^0,x^1) = [\sum_{i=1}^{N} s_i^0 (p_i^1/p_i^0)^{r/2}]^{1/r} [\sum_{j=1}^{N} s_j^1 (p_j^1/p_j^0)^{-r/2}]^{-1/r}. \tag{17}$$

It can be verified that when $r=2$, P_r defined by (17) coincides with the Fisher index P_2 defined earlier.

The following theorem relates c^r defined by (16) to the price index formula P_r defined by (17).

Theorem 5

(Diewert [1976; p.133]): Suppose $C(u,p) = uc^r(p)$ and the observed data (p^0,x^0) and (p^1,x^1) for periods 0 and 1 satisfy the cost minimization assumption (5). Then for any reference utility level $u>0$, we have

$$P_K(p^0,p^1,u) = c^r(p^1)/c^r(p^0) = P_r(p^0,p^1,x^0,x^1). \tag{18}$$

Thus $P_K(p^0,p^1,u)$ may be precisely determined by using the mechanical index number formula P_r defined by (17). Diewert [1976] calls a price index formula $P(p^0,p^1,x^0,x^1)$ **superlative** if it correctly evaluates the Konüs price index $P_K(p^0,p^1,u) \equiv C(u,p^0)/C(u,p^1)$ for some cost function C of the form $C(u,p)=uc(p)$ where the unit cost function c can provide a second order approximation to an arbitrary unit cost function. Equations (16) to (18) show that the price indexes P_r are superlative for each $r \neq 0$.

Theorem 5 provides a strong economic justification for the use of the indexes P_r in empirical applications. In particular, we have obtained an economic justification for the use of the Fisher index P_2.

However, Theorem 5 is subject to a serious defect: namely, if $C(u,p) = uc^r(p)$, then the preference function F^r that is dual to this cost function is homothetic; in particular $F^r(x)$ is linearly homogeneous in x. Theorem 6 below is not subject to this defect.

First, let us define the **translog cost** function C^0 by:

$$\ln C^0(u,p) \equiv a_0 + \sum_{i=1}^{N} a_i \ln p_i + 1/2 \sum_{i=1}^{N}\sum_{j=1}^{N} a_{ij} \ln p_i \ln p_j + a_{00} \ln u$$
$$+ \sum_{i=1}^{N} a_{0i} \ln p_i \ln u + 1/2\, a_{000} (\ln u)^2 \tag{19}$$

where the parameters a_{ij} satisfy the following restrictions.

$$\sum_{i=1}^{N} a_i = 1;\ a_{ij} = a_{ji};\ \sum_{j=1}^{N} a_{ij} = 0 \text{ for } i=1,\ldots,N;\ \sum_{i=1}^{N} a_{0i} = 0. \tag{20}$$

Define the translog price index P_0[12] by

$$P_0(p^0,p^1,x^0,x^1) \equiv \prod_{i=1}^{N} (p_i^1/p_i^0)^{(s_i^0 + s_i^1)/2} \tag{21}$$

Theorem 6

(Diewert [1976; p.122]: Suppose the consumer's cost function C equals the translog cost function C^0 defined by (19) and suppose the observed data (p^0,x^0) and (p^1,x^1) satisfy the cost minimization assumptions (5) where F is the utility function dual to C. Define $u^0 \equiv F(x^0)$, $u^1 \equiv F(x^1)$ and $u^* \equiv (u^0u^1)^{1/2}$. Then the true cost of living index $P_K(p^0,p^1,u^*)$

evaluated at the intermediate utility level u^* may be calculated by evaluating the translog price index P_0; i.e., we have

$$P_K(p^0,p^1,u^*) \equiv C^0(u^*,p^1)/C^0(u^*,p^0) = P_0(p^0,p^1,x^0,x^1). \tag{22}$$

We note that the translog cost function C^0 defined by (19) can provide a second order approximation to an arbitrary twice continuously differentiable cost function, so that Theorem 6 is not restricted to the homothetic case. Thus Theorem 6 provides a very strong economic justification for the use of the translog price index P_0 as an approximation to the true cost of living $P_K(p^0,p^1,u^*)$.

Theorems 4 and 5 together provided a strong justification for the use of the Fisher index P_2 while Theorem 6 justified the use of the translog index P_0. Which one should we use? We will return to this question in Section 4, but first, it is useful to study the problem of measuring changes in the consumer's welfare (as opposed to the problem of measuring changes in the levels of prices that he faces).

3. Single Household Welfare Indicators

A first approach to measuring changes in the consumer's welfare would be to use the Konüs cost-of-living index $P_K(p^0,p^1,u)$ as a deflator for the consumer's expenditure ratio between the two periods, $p^1 \cdot x^1/p^0 \cdot x^0$. Hence we define the *Pollak* [1971; p.64] *Implicit Quantity Index* $\tilde{Q}_K$ as

$$\tilde{Q}_K(p^0,p^1,x^0,x^1,u) \equiv p^1 \cdot x^1/p^0 \cdot x^0 P_K(p^0,p^1,u). \tag{23}$$

If we make our usual cost minimization assumption (5), and if we use definition (2), we find that we can rewrite (23) as

$$\tilde{Q}_K(p^0,p^1,x^0,x^1,u) = \frac{C(u^1,p^1)}{C(u,p^1)} \bullet \frac{C(u,p^0)}{C(u^0,p^0)}, \tag{24}$$

where as usual $u^0 \equiv F(x^0)$ and $u^1 \equiv F(x^1)$.

There are two natural choices for the reference utility level u; namely, u^0 and u^1. Inserting these choices into (24) leads to the following formulae:

$$\tilde{Q}_K(p^0,p^1,x^0,x^1,u^0) = C(u^1,p^1)/C(u^0,p^1); \tag{25}$$

$$\tilde{Q}_K(p^0,p^1,x^0,x^1,u^1) = C(u^1,p^0)/C(u^0,p^0). \tag{26}$$

Diewert [1981: p.170] shows that $\tilde{Q}_K$ defined by (24) has the correct ordinal properties if the reference utility level u is chosen to be any level between u^0 and u^1 (including u^0 and u^1 as well); i.e., if $u^0 \leq u \leq u^1$ with at least one strict inequality, then $\tilde{Q}_K(p^0,p^1,x^0,x^1,u) > 1$ while if $u^0 \geq u \geq u^1$ with at least one strict inequality, then

$$\tilde{Q}_K(p^0,p^1,x^0,x^1,u) < 1.$$

The special cases of (23) defined by (25) and (26) may be illustrated with reference to Figure 1. The index defined by (25) is equal to the ratio OD/OB while the index defined by (26) is equal to the ratio OC/OA.

The special cases (25) and (26) are also special cases of another class of quantity indexes. For $x^0 > 0_N$, $x^1 > 0_N$ and $p >> 0_N$, define the *Allen* [1949; p.199] *Quantity Index*[13] as

$$Q_A(x^0,x^1,p) \equiv C[F(x^1), p]/C[F(x^0),p]. \tag{27}$$

When $p=p^1$, (27) reduces to (25), and when $p=p^0$, (27) reduces to (26). The reader will also be able to verify that the implicit quantity index $\tilde{Q}_K$ defined by (24) is a product of two Allen quantity indexes.

The Pollak and Allen quantity indexes, $\tilde{Q}_K$ and Q_A, are studied in greater detail (and

bounds are derived) in Pollak [1971] and Diewert [1981]. We shall not dwell on their properties here since neither index is the most natural concept for a quantity index or a welfare indicator. If $x^1 = \lambda x^0$ for some $\lambda > 0$, it would be desirable if our quantity index took on the value λ. Neither $\tilde{Q}_K$ nor Q_A has this desirable homogeneity property. However, the *Malmquist* [1953; 232] *Quantity Index* Q_M does have this homogeneity property. In order to define Q_M, we must first define the **deflation** or distance function $D(u,x)$ that corresponds to the consumer's utility function F. For $u > 0$ and $x >> 0_N$, define

$$D(u,x) \equiv \max_k \{k: F(x/k) \geq u, k > 0\}. \tag{28}$$

Thus $D(u,x^1)$ is the deflation factor k_1 say that will just reduce the vector x^1 proportionately so that $F(x^1/k_1) = u$. If F satisfies Conditions I, then D will satisfy certain regularity conditions (Conditions III say) and a D satisfying these conditions will uniquely characterize F.[14]

For $u > 0$, $x^0 >> 0_N$, $x^1 >> 0_N$, define the *Malmquist Quantity Index* Q_M as

$$Q_M(x^0,x^1,u) \equiv D(u,x^1)/D(u,x^0). \tag{29}$$

In general, the Malmquist quantity index $Q_M(x^0,x^1,u)$ will depend on the reference indifference surface indexed by u. As usual, two natural choices for u are $u^0 \equiv F(x^0)$ and $u^1 \equiv F(x^1)$. Thus the *Laspeyres-Malmquist quantity index* is defined as

$$\begin{aligned} Q_M(x^0,x^1,u^0) &\equiv D(u^0,x^1)/D(u^0,x^0) \\ &= D(u^0,x^1) \end{aligned} \tag{30}$$

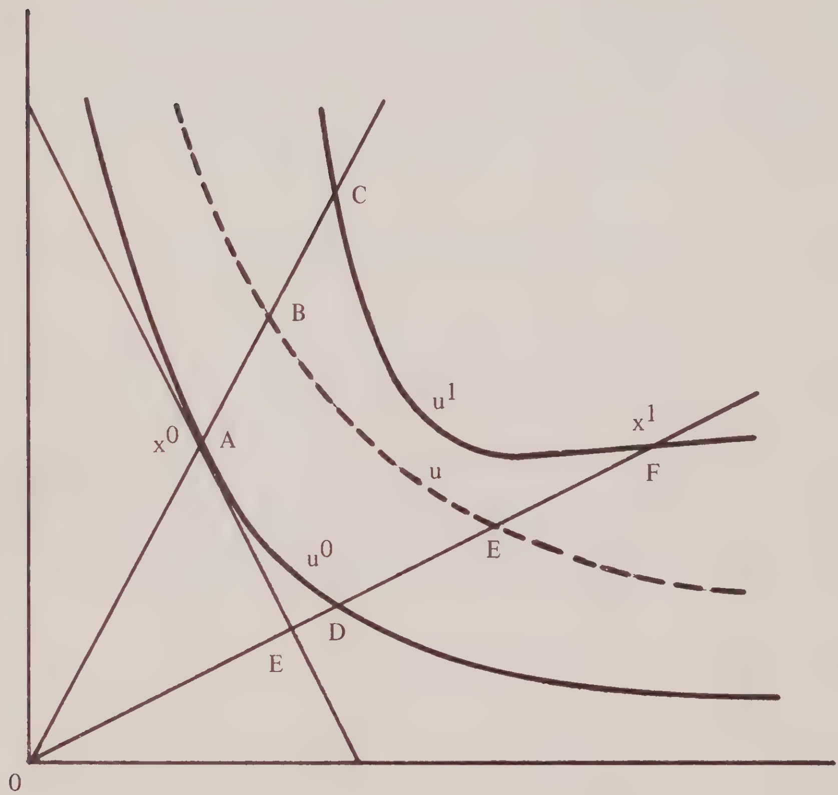
where (30) follows since $D(u^0,x^0) = 1$. The *Paasche-Malmquist quantity index* is defined as

$$\begin{aligned} Q_M(x^0,x^1,u^1) &\equiv D(u^1,x^1)/D(u^1,x^0) \\ &= 1/D(u^1,x^0) \end{aligned} \tag{31}$$

since $D(u^1,x^1) = 1$.

Geometric interpretations for the general Malmquist index (29) and the two special indexes (30) and (31) may be obtained from Figure 2 for the case of two goods. An observed quantity vector x^0 is the point A on the u^0 indifference curve while the other observed quantity vector x^1 is the point F on the u^1 indifference curve. The reference indifference curve indexed by u is indicated by the dashed indifference curve.

Figure 2.



The reader can confirm that $Q_M(x^0,x^1,u)$ defined by (29) is equal to [OF/0E]/[OA/OB], that $Q_M(x^0,x^1,u^0) = OF/OD$ and that $Q_M(x^0,x^1,u^1) = OC/OA$.

Note that the assumption of cost minimizing behaviour is **not** required in order to define the Malmquist quantity index.

From Figure 2, we see that as long as the reference indifference surface indexed by u remains between u^0 and u^1, the Malmquist index $Q_M(x^0,x^1,u)$ will correctly indicate whether welfare has increased or decreased going from x^0 to x^1. This property holds in general as the following result indicates.[15]

Theorem 7

(Diewert [1981; p.174]): If F satisfies Conditions I, $x^0 >> 0_N$, $x^1 >> 0_N$ and u is between $F(x^0)$ and $F(x^1)$, then (i) $Q_M(x^0,x^1,u) = 1$ if $F(x^0) = F(x^1)$, (ii) $Q_M(x^0,x^1,u) < 1$ if $F(x^0) > F(x^1)$, and (iii) $Q_M(x^0, x^1,u) < 1$ if $F(x^0) < F(x^1)$.

Comparing definitions (2), which defined P_K as a ratio of cost functions, with definition (29), which defined Q_M as a ratio of distance functions, it could be conjectured that Q_M will satisfy more or less the same mathematical properties as P_K, except that the role of prices and quantities is interchanged. This conjecture is correct. Thus we could write down a Q_M version of Theorem 1, where Q_M replaces P_K, x's replace p's, D replaces C, and Conditions III for distance functions D replace Conditions II for cost functions C. Similarly, Theorem 8 below is an exact analogue to Theorem 2.

Theorem 8

(Diewert [1981; p.175]): If F satisfies Conditions I and $x^0 >> 0_N$, $x^1 >> 0_N$, $u>0$, then

$$\min_i\{x_i^1/x_i^0\} \le Q_M(x^0,x^1,u) \le \max_i \{x_i^1/x_i^0\}.$$

The above theorem provides a start to the problem of providing observable points for the essentially unobservable index Q_M. The above theorem did not require the assumption of cost minimizing behaviour on the part of the consumer. The following theorem does.

Theorem 9

(Malmquist [1953; p.231]): Suppose F satisfies Conditions I and (5) holds. Define $u^0 \equiv F(x^0)$ and $u^1 \equiv F(x^1)$. Then

$$Q_M(x^0,x^1,u^0) \leq p^0 \cdot x^1/p^0 \cdot x^0 \equiv Q_L(p^0,p^1,x^0,x^1), \text{ and} \tag{33}$$

$$Q_M(x^0,x^1,u^1) \geq p^1 \cdot x^1/p^1 \cdot x^0 \equiv Q_P(p^0,p^1,x^0,x^1). \tag{34}$$

Leontief [1936; pp.58-59] illustrated the above bounds in the two good case. Note that the right-hand side of (33) defines the *Laspeyres quantity index* Q_L and the left-hand side of (33) is the Laspeyres-Malmquist quantity index defined earlier by (30). Thus Q_L is an observable upper bound to the essentially unobservable Malmquist index $Q_M(x^0,x^1,u^0)$. In Figure 2, $Q_M(x^0,x^1,u^0) = 0F/0D \leq 0F/0G = p^0 \cdot x^1/p^0 \cdot x^0 = Q_L$.

The right-hand side of (34) defines the *Paasche quantity index* Q_P and the left-hand side of (34) is the Paasche-Malmquist quantity index defined earlier by (31). Thus Q_P is an observable lower bound to $Q_M(x^0,x^1,u^1)$.

Corollary

$$\min_i \{x_i^1/x_i^0\} \leq Q_M(x^0,x^1,F(x^0)) \leq Q_L \equiv p^0 \cdot x^1/p^0 \cdot x^0; \tag{35}$$

$$p^1 \cdot x^1/p^1 \cdot x^0 \equiv Q_P \leq Q_M(x^0,x^1,F(x^1)) \leq \max_i \{x_i^1/x_i^0\} \tag{36}$$

Note that $Q_M(p^0,p^1,x^0,x^1)\ P_P(p^0,p^1,x^0,x^1) = p^1 \cdot x^1/p^0 \cdot x^0 = Q_P(p^0,p^1,x^0,x^1)\ P_L(p^0,p^1,x^0,x^1)$. Thus if the Paasche and Laspeyres price indexes P_P and P_L are numerically close, then the Paasche and Laspeyres price indexes P_P and P_L are numerically close, then the Paasche and Laspeyres quantity indexes Q_P and Q_L will also be close. Thus the upper bound in (35) will often be close to the lower bound in (36). Hence the following theorem is extremely useful.

Theorem 10

(Diewert [1981; p.176]): Suppose F satisfies Conditions I and the cost minimization assumption (5) holds. Then there exists a reference utility level u^* between $u^0 \equiv F(x^0)$ and $u^1 \equiv F(x^1)$ such that the Malmquist quantity index $Q_M(x^0,x^1,u^*)$ for this reference utility level u^* lies between Q_L and Q_P.

Theorem 10 is a counterpart to Theorem 4 (and in fact is proved in the same manner). Thus if Q_L and Q_P are "close" to each other, we may take an average of them (such as Irving Fisher's ideal quantity index $Q_2 \equiv (Q_L Q_P)^{1/2}$) and obtain a "close" approximation to the unobservable Malmquist quantity index $Q_M(x^0,x^1,u^*)$ where u^* is some utility level between u^0 and u^1.

Note that Theorem 10 made no assumptions about the shape of the consumer's indifference surfaces (other than our usual general regularity conditions on the consumer's utility function F). However our choice of the reference utility level u^* was somewhat limited.

In order to make further progress, it is necessary to make specific functional form assumptions. Thus let $f^r(x)$ be the quadratic mean of order r utility function for $r \neq 0$, defined in a manner analogous to the definition of $c^r(p)$ (recall (16)) and define the quadratic mean of order r mechanical quantity index formula Q_r by

$$Q_r(p^0,p^1,x^0,x^1) \equiv P_r(x^0,x^1,p^0,p^1) \qquad (37)$$

where P_r was defined by (17). Note that we have interchanged the role of prices and quantities in the right-hand side of (37). It may be verified that when $r = 2$, Q_2 defined by (37) reduces to the Fisher quantity index, $(Q_P Q_L)^{1/2}$.

The following theorem is the quantity counterpart to Theorem 5.

Theorem 11

(Diewert [1976; p.132]): Suppose $F = f^r$ for some $r \neq 0$ and the data (p^0,x^0), (p^1,x^1) satisfy the cost minimization assumption (5). Then for any reference utility level $u > 0$, we have

$$Q_M(x^0,x^1,u) = f^r(x^1)/f^r(x^0) = Q_r(p^0,p^1,x^0,x^1). \tag{38}$$

Thus $Q_M(x^0,x^1,u)$ may be precisely determined by using the mechanical index number formula Q_r defined by (37), provided that $F = f^r$. Since f^r can provide a second order approximation to an arbitrary twice continuously differentiable homogeneous utility function, Diewert [1976] calls Q_r a **superlative** quantity index number formula.

Although Theorem 11 provides a strong economic justification for the use of the superlative indexes Q_r, the result is subject to the same defect that occurred in Theorem 5; namely, the preferences that correspond to f^r are homothetic. Theorem 12 below is not subject to this limitation.

First, we define the translog distance function $D^0(u,x)$ by setting $\ln D^0(u,x)$ equal to the right-hand side of (19) where x_i replaces p_i. Define the *translog quantity index* Q_0 by

$$\ln Q_0(p^0,p^1,x^0,x^1) = \sum_{i=1}^{N} \frac{1}{2}\left\{\frac{p_i^0x_i^0}{p^0\cdot x^0} + \frac{p_i^1x_i^1}{p^1\cdot x^1}\right\} \ln(x_i^1/x_i^0). \tag{39}$$

Theorem 12

(Diewert [1976; p.123]): Suppose the consumer's distance function D equals the translog distance function D^0 defined above. Let F be the corresponding utility function and C the corresponding cost function. Let the observed data (p^0,x^0), (p^1,x^1) satisfy the cost minimization assumptions (5). Define $u^0 \equiv F(x^0)$, $u^1 \equiv F(x^1)$ and $u^* \equiv (u^0u^1)^{1/2}$. Then the Malmquist quantity index $Q_M(x^0,x^1,u^*)$ evaluated at the intermediate utility level u^* is precisely equal to the translog quantity index Q_0; i.e.,

$$Q_M(x^0,x^1,u^*) = Q_0(p^0,p^1,x^0,x^1).$$

Since the translog distance function D^0 can provide a second order approximation to an arbitrary twice continuously differentiable distance function, Theorem 12 provides a very strong economic justification for the use of the translog quantity index Q_0 as an approximation to the Malmquist quantity index $Q_M(x^0,x^1,u^*)$.

4. Fixed Base versus Chain Indexes

In Section 2, we found a family of mechanical index number formulae, $P_r(p^0,p^1,x^0,x^1)$ for each number r, which had *a priori* good properties from an economic point of view. To each such P_r, there corresponds an implicit quantity index $\tilde{Q}_r$ that may be defined as follows:

$$\tilde{Q}_r(p^0,p^1,x^0,x^1) \equiv p^1 \cdot x^1 / p^0 \cdot x^0 \; P_r(p^0,p^1,x^0,x^1). \tag{40}$$

In Section 3, we found a family of mechanical quantity indexes, $Q_r(p^0,p^1,x^0,x^1)$ for each number r, which had *a priori* good properties from an economic point of view. To each such Q_r, there corresponds an implicit price index $\tilde{P}_r$ defined by:

$$\tilde{P}_r(p^0,p^1,x^0,x^1) \equiv p^1 \cdot x^1 / p^0 \cdot x^0 \; Q_r(p^0,p^1,x^0,x^1). \tag{41}$$

Our problem now is that we have too many index number formulae that have desirable properties. Which price index formula should we choose from the double infinity of candidates of the form P_r or $\tilde{P}_s$ for some r or s? The following theorem leads to an answer to this question.

Theorem 13

(Diewert [1978]): The functions $P_r(p^0,p^1,x^0,x^1)$ and $\tilde{P}_s(p^0,p^1,x^0,x^1)$ differentially approximate each other to the second order around any point where $p^0 = p^1 >> 0_N$ and $x^0 = x^1 >> 0_N$. A similar statement holds for the quantity indexes Q_s and $\tilde{Q}_r$.

Thus we have for all r and s;

$$P_r(p^0,p^0,x^0,x^0) = \tilde{P}_s(p^0,p^0,x^0,x^0) \tag{42}$$

$$\nabla P_r(p^0,p^0,x^0,x^0) = \nabla\tilde{P}_s(p^0,p^0,x^0,x^0) \text{ and} \tag{43}$$

$$\nabla^2 P_r(p^0,p^0,x^0,x^0) = \nabla^2\tilde{P}_s(p^0,p^0,x^0,x^0) \tag{44}$$

where ∇P_r stands for the 4N dimensional vector of first order partial derivatives of the function P_r and $\nabla^2 P_r$ is the 4N by 4N matrix of second order partials of P_r.

Hence for "normal" time series data, the indexes $P_r(p^0,p^1,x^0,x^1)$ and $\tilde{P}_s(p^0,p^1,x^0,x^1)$ **will all give the same answer** to a very high degree of approximation. Some empirical evidence on this issue is presented in Diewert [1978; 1983b] and in Généreux [1983]. Thus the answer to the question posed before Theorem 13 is: it does not matter very much which formula we use.

Diewert [1978] also shows that the Paasche and Laspeyres indexes that figured prominently in sections 2 and 3 approximate the superlative indexes to the first order, i.e., for $p^0 >> 0_N$ and $x^0 >> 0_N$, we have

$$P_r(p^0,p^0,x^0,x^0) = P_L(p^0,p^0,x^0,x^0) = P_P(p^0,p^0,x^0,x^0) \text{ and} \tag{45}$$

$$\nabla P_r(p^0,p^0,x^0,x^0) = \nabla P_L(p^0,p^0,x^0,x^0) = \nabla P_P(p^0,p^0,x^0,x^0) \tag{46}$$

but we do not obtain the equality of the second order partial derivatives as we did in (44).

It is worth emphasizing that the above results hold **without** the assumption of optimizing behaviour on the part of consumers; i.e., they are theorems in numerical analysis rather then economics.[16]

Theorems 4 and 10 above showed that it is very useful to have the Paasche and Laspeyres indexes close to each other since this will lead to a very close approximation to the Konüs

cost-of-living index $P_K(p^0,p^1,u^*)$ and to the Malmquist quantity index $Q_M(x^0,x^1,u^*)$. This implies that we should use the **chain** principle for constructing indexes rather than the **fixed base** principle. The difference between the two principles may be illustrated for the case of three observations. In the fixed base principle, given a mechanical index number formula P, the aggregate level of prices for the three periods would be

$$1,\ P(p^0,p^1,x^0,x^1) \text{ and } P(p^0,p^2,x^0,x^2). \tag{47}$$

Using the chain principle, the aggregate price level for the three periods would be

$$1,\ P(p^0,p^1,x^0,x^1) \text{ and } P(p^0,p^1,x^0,x^1)P(p^1,p^2,x^1,x^2).$$

Using the chain principle, variations between prices and quantities tend to be smaller and hence the Paasche, Laspeyres and all of the superlative indexes will approximate each other much more closely than is the case when we use a fixed base. This is borne out in empirical computations.[17]

If we had to choose a particular index number formula for a price index from the families P_r and $\tilde{P}_s$, we could perhaps narrow the choice to three indexes: (i) P_0, because of Theorem 6, (ii) $\tilde{P}_0$, because of Theorem 12, or (iii) P_2, because it lies between the Paasche and Laspeyres bounds.[18]

Of course, there is a high **empirical** cost associated with the use of a superlative index number formula as opposed to using a simple fixed based Laspeyres index of the form $p^t \cdot x^0 / p^0 \cdot x^0$. The advantage of this latter index is that it requires quantity information for only the base period. Its disadvantage is that it will probably[19] not approximate the consumer's true cost-of-living index $P_K(p^0,p^t,u^0)$ very closely as p^t moves further away from p^0.

Having discussed price and welfare indexes for an individual consumer or household at great length, it is time now to turn to the construction of group indexes.

5. The Democratic Consumer Price Index

Assume that there are H distinct households, where household h has utility function $F^h(x)$ satisfying Conditions I noted in Section 2 with a dual cost function $C^h(u_h,p)$.

There are many possible ways for constructing a group cost-of-living index. Our first method will be to construct a simple average of the individual household Konüs cost-of-living indexes $C^h(u_h,p^1)/C^h(u_h,p^0)$. This is what Prais [1959] and Muellbauer [1974] call a *democratic price index*. Letting $u \equiv (u_1,\ldots,u_H)^T$ denote a vector of reference utility levels, we define the *democratic cost-of-living index* P_D as

$$P_D(p^0,p^1,u) \equiv \sum_{h=1}^{H} \frac{1}{H} \frac{C^h(u_h,p^1)}{C^h(u_h,p^0)}. \tag{49}$$

Let $p^0 >> 0_N$ and $p^1 >> 0_N$ be the period 0 and 1 price vectors as usual, let $x_h^t \equiv (x_{1h}^t,x_{2h}^t,\ldots,x_{Nh}^t)^T$ be consumer h's observed quantity vector for period t, $t=0$, 1 and for $h = 1,\ldots, H$, let $u_h^0 \equiv F^h(x_h^0)$ and $u_h^1 \equiv F^h(x_h^1)$ be the period 0 and 1 utility levels attained by household h, and define the utility vectors $u^1 \equiv (u_1^1,\ldots,u_H^1)^T$ and $u^0 \equiv (u_1^0,\ldots,u_H^0)^T$. Assume cost minimizing behaviour for each household during both periods; i.e.,

$$p^0 \cdot x_h^0 = C^h(u_h^0,p^0) = C^h[F^h(x_h^0),p^0] \qquad \text{for } h=1,\ldots,H \tag{50}$$

and

$$p^1 \cdot x_h^1 = C^h(u_h^1,p^1) = C^h[F^h(x_h^1),p^1] \qquad \text{for } h=1,\ldots,H.$$

Then we may immediately apply Theorems 2 and 3 in order to obtain the following bounds on the Laspeyres-Democratic index $P_D(p^0,p^1,u^0)$ and on the Paasche-Democratic index $P_D(p^0,p^1,u^1)$:

$$\min_i\{p_i^1/p_i^0\} \leq P_D(p^0,p^1,u^0) \leq \sum_{h=1}^{H} \frac{1}{H} \frac{p^1 \cdot x_h^0}{p^0 \cdot x_h^0}; \tag{51}$$

$$\sum_{h=1}^{H} \frac{1}{H} \frac{p^1 \cdot x_h^1}{p^0 \cdot x_h^1} \leq P_D(p^0,p^1,u^1) \leq \max_i\{p_i^1/p_i^0\}. \tag{52}$$

Note that the right-hand side of (51) is an arithmetic average of the individual household Laspeyres price indexes, $\Sigma_h P_L(p^0,p^1,x_h^0,x_h^1)/H$, which we denote by $\bar{P}_L$. The left-hand side of (52) is an arithmetic average of the individual Paasche indexes, $\Sigma_h P_P(p^0,p^1,x_h^0,x_h^1)/H$, which we denote by $\bar{P}_P$.

The following theorem provides a group counterpart to Theorem 4 (and in fact is proven in the same manner).

Theorem 14

Let each consumer's utility function F^h satisfy Conditions I. Suppose that the observed data for periods 0 and 1,(p^0,x_h^0) and (p^1,x_h^1) for $h=1,\ldots H$, satisfy the cost minimization assumption (50). Then there exists a reference utility vector $u^* \equiv (u_1^*,\ldots,u_H^*)$ such that each component u_h^* lies between u_h^0 and u_h^1 and the Prais-Muellbauer group cost-of-living index evaluated at this reference utility vector $P_D(p^0,p^1,u^*)$, lies between $\bar{P}_L$ and $\bar{P}_P$.

For typical data, we would expect $\bar{P}_L$ and $\bar{P}_L$ to be very close to each other (closer than the individual indexes P_L^h and P_P^h since $\bar{P}_L$ and $\bar{P}_P$ are **averages** of the individual P^h, and P_P^h), so $P_D(p^0,p^1,u^*)$ may be very closely approximated by either $\bar{P}_L$ or $\bar{P}_P$.

The practical difficulty with using the democratic price index $P_D(p^0,p^1,u^*)$ as a general measure of inflation between periods 0 and 1 is that $\bar{P}_L$ and $\bar{P}_P$ can only be constructed if we have individual household data. The group index defined in the following section has bounds that can be constructed from aggregate data.

6. The Plutocratic Cost-of-Living Index

Pollak [1981; p.328] defines what he calls a Scitovsky-Laspeyres cost-of-living index for a group of consumers as the ratio of the total expenditure required to enable each household to attain its reference indifference curve at period 1 prices to that required at period 0 prices. This same concept of a group cost-of-living index was suggested by Prais [1959] in less precise language. Prais referred to his concept as a *plutocratic price index*. Making the same assumptions and using the same notation as in the previous section, this Prais-Pollak *plutocratic cost-of-living index* may be defined as:

$$P_{PP}(p^0,p^1,u) \equiv \sum_{h=1}^{H} C^h(u_h,p^1) \Big/ \sum_{h=1}^{H} C^h(u_h,p^0) \tag{53}$$

$$= \sum_{h=1}^{H} s^h(u,p^0)\, C^h(u_h,p^1)/C^h(u_h,p^0) \tag{54}$$

where the consumer h share of total expenditure at reference utility levels u prices p is defined as

$$s^h(u,p) \equiv C^h(u_h,p)\Big/ \sum_{h=1}^{H} C^h(u_h,p) \qquad \text{for } h=1,\ldots,H. \tag{55}$$

The *Laspeyres-Pollak cost-of-living index* is defined as $P_{PP}(p^0,p^1,u^0)$ while the *Paasche-Pollak cost-of-living index* is defined as $P_{PP}(p^0,p^1,u^1)$. Note that each household is given the same weight $1/H$ in the democratic price index defined by (49), while household h is given the expenditure share weight $s^h(u,p^0)$ in the plutocratic price index defined by (54).

It is convenient to define aggregate consumption vectors $\bar{x}^t$ as follows:

$$\bar{x}^0 \equiv \sum_{h=1}^{H} x_h^0 ; \qquad \bar{x}^1 \equiv \sum_{h=1}^{H} x_h^1 . \tag{56}$$

Armed with the above definitions, we may now prove the usual Paasche and Laspeyres bounding theorem.

Theorem 15

Suppose each consumer's utility function F^h satisfies Conditions I and suppose the cost minimization assumptions (5) hold. Then

$$\min_i\{p_i^1/p_i^0\} \le P_{PP}(p^0,p^1,u^0) \le p^1 \cdot \bar{x}^0/p^0 \cdot \bar{x}^0 \equiv P_L \tag{57}$$

$$P_P \equiv p^1 \cdot \bar{x}^1/p^0 \cdot \bar{x}^1 \le P_{PP}(p^0,p^1,u^1) \le \max_i\{p_i^1/p_i^0\}. \tag{58}$$

Note that P_L in (57) is the usual Laspeyres price index involving the **aggregate** base period consumption vector $\bar{x}^0$, while P_P in (58) is the usual Paasche price index involving the **aggregate** period 1 consumption vector $\bar{x}^1$. The consumer price index constructed by statistical agencies is usually an approximation to P_L. All that is required to construct P_L is: the current vector of prices p^1, the base price vector p^0, and the base period aggregate consumption vector $\bar{x}^0 \equiv \Sigma_h x_h^0$.

Usually, the upper bound P_L in (57) will be close to $P_{PP}(p^0,p^1,u^0)$ while the lower bound P_P in (58) will be close to $P_{PP}(p^0,p^1,u^1)$ and to P_L.

Theorem 16

Under the conditions of Theorem 15, there exists a reference utility vector $u^* \equiv (u_1^*,\ldots,u_H^*)$ such that each component u_h^* lies between u_h^0 and u_h^1 and the Prais-Pollak group cost-of-living index evaluated at this reference utility vector, $P_{PP}(p^0,p^1,u^*)$, lies between $P_L \equiv p^1 \cdot \bar{x}^0/p^0 \cdot \bar{x}^0$ and $P_P \equiv p^1 \cdot \bar{x}^1/p^0 \cdot \bar{x}^1$.

As was the case for the Prais-Muellbauer index P_D, we would expect P_L and P_P to be very close to each other provided that p^0 and p^1 are not "too" different, and hence $P_{PP}(p^0,p^1,u^*)$ may be closely approximated by either P_L or P_P (or say by the Fisher ideal $P_2 \equiv P_L P_P{}^{1/2}$) under these circumstances.

We have noted that the bounds for the Pollak index, P_L and P_P, may be computed provided only that we have aggregate data on prices and quantities, while the bounds for the Muellbauer index, $\overline{P}_L$ and $\overline{P}_P$, require detailed household data for their computation. Are there other important differences between the two concepts of the group price index?

An important conceptual difference emerges if we compare the formula (49) for P_D and the formula (54) for P_{PP}. In (49), each household's individual Konüs cost-of-living index, $C^h(u_h,p^1)/C^h(u_h,p^0)$, is given the same weight, $1/H$. In (54), when $u = u^0$, household h's Laspeyres-Konüs cost-of-living index, $C^h(u_h^0,p^1)/C^h(u_h^0,p^0)$, is given the weight $s^h(u^0,p^0)$, which is household h's share of total expenditure during period 0. Thus high expenditure households will tend to get weighted more heavily than low expenditure households in the construction of the plutocratic index. This is why Prais and Muellbauer call the index (49) a democratic index, since it weights each household's cost-of-living index equally.

Which concept of a group price index should be used as a measure of average inflation between periods 0 and 1? Unless we have quantity data by household class, we cannot evaluate the bounds for the democratic index, so in this case we are stuck with the plutocratic index. If we do have detailed data by household class, then instead of calculating bounds for the democratic index, we should calculate the usual Paasche and Laspeyres bounds for the **individual** households true cost-of-living indexes, $C^h(u_h,p^1)/C^h(u_h,p^0)$.

Deaton and Muellbauer [1980; p.178] note that the simplification of working with a single price index can be very dangerous. They cite the Great Bengal Famine of 1943 when between three and five million people died of starvation. An average price index was not very relevant to the problems of low-income households under those circumstances.

Statistics Canada [1982; p.89] has taken the first step in the direction of providing consumer price indexes that are household specific in that they now construct an experimental CPI for low-income families. They are to be commended for this effort and they should be given additional resources to construct additional price indexes by household class.

We turn now to a brief discussion of the mirror image to the problem of constructing group price indexes - the problem of constructing aggregate welfare indexes.

7. Social Welfare Indexes

In Section 3, we found that the Malmquist quantity index, $Q_M(x^0,x^1,u) \equiv D(u,x^1)/D(u,x^0)$, was a very satisfactory quantity index or welfare indicator for a single household. Let us suppose that household h's preferences may be represented by the deflation function $D^h(u_h,x_h)$ that is dual to the utility function $F^h(x_h)$, where F^h satisfies Conditions I as usual.

Define the N by H matrix of period 0 (1) consumer choices by X^0 (X^1): i.e.,

$$X^0 \equiv [x_1^0,x_2^0,\ldots,x_H^0]; \qquad X^1 \equiv [x_1^1, x_2^1,\ldots,x_H^1] \tag{59}$$

where $x_h^t \equiv [x_{1h}^t, x_{2h}^t,\ldots,x_{Nh}^t]^T$ is consumer h's observed consumption vector during period t for t=0, 1 and h= 1, 2,...,H.

It is natural to try and aggregate individual welfare changes into a single scalar measure of overall welfare change. A simple *social index of welfare change* that respects individual preferences is

$$W(X^0,X^1,u) \equiv \sum_{h=1}^{H} \beta_h(u)\, D^h(u_h,x_h^1)/D(u_h,x_h^0) \tag{60}$$

where $u \equiv (u_1,u_2,\ldots,u_H)^T$ is a vector of individual reference utility levels and the weight function β_h satisfies the following restriction:

$$\sum_{h=1}^{H} \beta_h(u) = 1 \qquad \text{for any vector of reference utilities u.} \tag{61}$$

The reader will note that our indicator of social welfare change defined by (60) is a quantity counterpart to the Pollak cost-of-living index defined by (54) (which is an indicator of price change).

Our reason for imposing the restriction (61) is the following: if $x_h^1 = x_h^0$ for $h=1,...,H$ so that there is no change in consumption between periods 0 and 1, then we want our indicator of social welfare change (or our aggregate quantity index) to indicate that there has been no change; i.e., we want $W(X^0,X^0,u)=1$. Hence we must have (61).

In general, the weights $\beta_h(u)$ do not have to be positive or even non-negative. However, if the weights are non-negative, then we obtain the following bounds for W applying Theorem 8:

$$\min_{i,h}\{x_{ih}^1/x_{ih}^0\} \leq W(X^0,X^1,u) \leq \max_{i,h}\{x_{ih}^1/x_{ih}^0\}. \tag{62}$$

As usual, we define the *Laspeyres Social Welfare Indicator* $W(X^0,X^1,u^0)$ by selecting our reference utility vector to be $u^0 \equiv [F^1(x_1^0),...,F^H(x_H^0)] \equiv [u_1^0,...,u_H^0]$ and the *Paasche Social Welfare indicator* $W(X^0,X^1,u^1)$ by selecting the reference utility vector to be $u^1 \equiv [u_1^1,u_2^1,...,u_H^1]$.

The following theorem provides a social welfare counterpart to Theorem 9.

Theorem 17

Let each household utility function F^h satisfy Conditions I and suppose the cost minimization assumptions (50) hold. Define the utility weights $\beta_h^0 \equiv \beta(u^0)$ and the utility weights $\beta_h^1 \equiv \beta_h(u^1)$ for $h=1,2,...,H$. Then if the weights β_h^0 and β_h^1 are non-negative,

$$W(X^0,X^1,u^0) \leq \sum_{h=1}^{H} \beta_h^0\, p^0\cdot x_h^1/p^0\cdot x_h^0 \equiv \sum_{h=1}^{H} \beta_h^0\, Q_L^h \tag{63}$$

and

$$\sum_{h=1}^{H} \beta_h^1\, Q_P^h \equiv \sum_{h=1}^{H} \beta_h^1\, p^1\cdot x_h^1/p^1\cdot x_h^0 \leq W(X^0,X^1,u^1). \tag{64}$$

Thus the Laspeyres social welfare indicator $W(X^0,X^1,u^0)$ is bounded from above by a weighted average of the individual household Laspeyres quantity indexes Q_L^h, and the Paasche social welfare indicator $W(X^0,X^1,u^1)$ is bounded from below by a weighted average of the individual household Paasche quantity indexes Q_P^h.

The following theorem provides a social welfare analogue to Theorem 10.

Theorem 18

Suppose that the hypotheses of Theorem 17 are satisfied. Suppose also that the household weighting functions $\beta_h(u)$ are continuous as u varies linearly between u^0 and u^1. Then there exists a reference utility vector $u^* = (1-\lambda^*)u^0 + \lambda^* u^1$ for some λ^* between 0 and 1 such that the social welfare change indicator $W(X^0,X^1,u^*)$ lies between $\Sigma_{h=1}^H \beta_h^0 Q_L^h$ (the upper bound in (63)) and $\Sigma_{h=1}^H \beta_h^1 Q_P^h$ the lower bound in (64)).

Corollary

Suppose $\beta_h^0 = p^0 \cdot x_h^0 / \Sigma_{k=1}^H p^0 \cdot x_k^1$ (the share of household h in base period expenditure) and $\beta_h^0 = p^1 \cdot x_h^0 / \Sigma_{k=1}^H p^1 \cdot x_k^0$ (the share of household h in expenditure using period 1 prices and period 0 quantities) for $h=1,\ldots,H$. Then the social welfare change indicator $W(X^0,X^1,u^*)$ defined in the theorem lies between the aggregate Laspeyres and Paasche quantity indexes, $p^0 \cdot (\Sigma x_h^1)/p^0 \cdot (\Sigma x_h^0)$ and $p^1 \cdot (\Sigma x_h^1)/p^1 \cdot (\Sigma x_h^0)$, respectively.

Thus if the aggregate Paasche and Laspeyres quantity indexes are close to each other, an average of them such as $Q_2 \equiv (Q_L Q_P)^{1/2}$ (Fisher's ideal quantity index), will yield a close approximation to the welfare change indicator $W(X^0,X^1,u^*)$ described in the above corollary. Thus we have provided a justification of sorts for Pigou's [1920; p.84] cautious recommendation of the Fisher quantity index as an indicator of aggregate welfare change.

However, I do not think that the corollary to Theorem 18 should be taken too seriously. We need very special welfare weights $\beta_h(u)$ in order to obtain the corollary. These special weights need not correspond to anybody's idea of a just society.

The paragraph above illustrates a problem with the social welfare approach: it is difficult to come to a consensus on what the weights $\beta_h(u)$ should be. Hence perhaps we should concentrate on the calculation of welfare change by household class rather than attempting to construct somewhat arbitrary measures of aggregate welfare change. For alternative attempts to construct aggregate cost-of-living indexes and aggregate measures of welfare change, see Blackorby and Donaldson [1983] and Jorgenson and Slesnick [1983].

We leave the world of group indexes in the following sections in order to focus on some special problems that are associated with single household price and quantity indexes.

8. The Theory of Subindexes

Pollak [1975a] noted that in many instances, economists are interested in *subindexes* of the cost-of-living index; i.e., in indexes that do not cover the whole spectrum of consumer goods, but only selected subsets. He notes several interesting classes of subindexes: (1) a food index say in the context of a complete cost-of-living index defined over all consumer goods, (2) a one-period index in the context of a multi-period world, (3) a consumption goods subindex in the context of a consumer choice model that included not only the consumption decision, but also the labour supply decision, (4) a consumption goods subindex in the context of a model where the consumer has preferences defined over not only consumer goods but also environmental variables (such as pollution) and also public goods (such as roads and parks).

Pollak's discussion summarized above indicated that the concept of a subindex in the cost-of-living index is not without applications. We must now face up to two problems: (i) how do we define a subindex rigorously, and (ii) how may we combine subindexes in order to form an approximation to the true overall cost-of-living index.[20]

Some new notation and a new concept are required. As usual, think of x as being the consumer's overall consumption vector. We now partition the vector x into M subvectors (of varying dimension) which we denote by $(x_1,x_2,\ldots,x_M) \equiv x$ (so x_m is the mth subvector). Partition the overall price vector p in an analogous manner. We may now define

Pollak's [1969][21] *conditional expenditure or cost function* for the mth subgroup of goods as:

$$C^m(u,p_m,x_1,x_2,\ldots,x_{m-1},x_{m+1},x_{m+2},\ldots,x_M) \tag{65}$$

$$\equiv \min_{x_m}\{p_m \cdot x_m : F(x) \geq u\} \qquad m=1,2,\ldots,M,$$

where u is the consumer's reference utility level, F is his overall utility function satisfying the usual conditions, and the overall consumption vector is $x \equiv (x_1,x_2,\ldots,x_{m-1},x_m,x_{m+1},\ldots,x_M)$. Thus C^m defined by (65) is the minimum group m cost of achieving utility level u, given that the consumer has available x_1 units of group 1 goods, x_2 units of group 2 goods,...,x_{m-1} units of group m-1 goods, x_{m+1} units of group m+1 goods,..., and x_M units of group M goods. In order to save space, we shall write C^m defined by (65) as $C^m(u,p_m,x)$ where x is the entire consumption vector, but it should be understood that $C^m(u, p_m, x)$ is constant with respect to variations in the x_m components of x; i.e., $C^m(u,p_m,x)$ does not depend on the x_m components of X.

The regularity properties of C^m with respect to the vector of price variables p_m are exactly the same properties that were given in Conditions II for the vector of price variables p.[22] Thus we may define Pollak's [1975; 147] *Subindex of the Cost of Living Index* for Group m goods as:

$$P^m(p_m^0,p_m^1,u,x) \equiv C^m(u,p_m^1,x)/C^m(u,p_m^0,x); \qquad m=1,\ldots,M, \tag{66}$$

where $p^0 \equiv (p_1^0,p_2^0,\ldots,p_M^0)$ is the base period price vector (remember each p_m^0 is a vector pertaining to group m goods), $p^1 \equiv (p_1^1,p_2^1,\ldots,p_M^1)$ is the period 1 price vector, u is a reference utility level, and $x \equiv (x_1,x_2,\ldots x_M)$ is a reference quantity vector.

Since Theorem 2 depended only on the regularity properties of C(u,p) with respect to the price vector p and since $C^m(u, p_m,x)$ satisfies these same regularity properties, it can

be seen that P^m satisfies the usual Lerner-Joseph-Samuelson bounds for any reference vector (u,x):

$$\min_i \{p^1_{mi}/p^0_{mi}\} \leq P^m(p^0_m,p^1_m,u,x) \leq \max_i \{p^1_{mi}/p^0_{mi}\}$$

where the index i runs through the components of the group m price vectors.

As usual, we can pick out particular reference vectors of interest. Define the *Laspeyres-Pollak Group* m *Subindex* as $P^m(p^0_m,p^1_m,u^0,x^0)$ where x^0 is the consumer's period 0 choice vector and $u^0 \equiv F(x^0)$ is his period 0 utility level. Define the *Paasche-Pollak Group* m *Subindex* as $P^m(p^0_m,p^1_m,u^1,x^1)$ where x^1 is the consumer's period 1 choice vector and $u^1 \equiv F(x^1)$ is his period 1 utility level. Assuming optimizing behaviour during the two periods, we may derive the following subindex counterpart to Theorem 3.

Theorem 19

Assuming (5), the mth subindex P^m defined by (66) satisfies the following inequalities when $(u,x) = (u^0,x^0)$ and (u^1,x^1) respectively:

$$P^m(p^0_m,p^1_m,u^0,x^0) \leq P^m_L \equiv p^1_m \cdot x^0_m/p^0_m \cdot x^0_m; \qquad m=1,\ldots,M \text{ and} \tag{67}$$

$$P^m(p^0_m,p^1_m,u^1,x^1) \geq P^m_P \equiv p^1_m \cdot x^1_m/p^0_m \cdot x^1_m; \qquad m=1,\ldots,M. \tag{68}$$

Thus we obtain the usual Laspeyres and Paasche bounds for the subindexes, and hence we may obtain the following adaptation of Theorem 4.

Theorem 20

Assume that F satisfies Conditions I, and assume that (5) holds; i.e., that there is overall utility maximizing behaviour during the two periods. Then there exists a reference utility-consumption vector $(u^{m*},x^{m*}) \equiv (1-\lambda^*_m)\,(u^0,x^0) + \lambda^*_m(u^1,x^1)$ where $0<\lambda^*_m<1$ and $u^0 \equiv F(x^0)$, $u^1 \equiv F(x^1)$ such that the mth Pollak subindex evaluated at this reference vector, $P^m(p^0_m,p^1_m,u^{m*},x^{m*})$,[23] lies between the group m Laspeyres index P^m_L defined in (67) and

the group m Paasche index P_P^m defined in (68).

Thus under normal circumstances when P_L^m is close to P_P^m, we may obtain a rather good estimate of the subindex $P^m(p_m^0,p_m^1,u^{m*},x^{m*})$.

How may we combine the subindexes in order to obtain an estimate of the overall cost-of-living $P_K(p^0,p^1,u)$?

Consider the Laspeyres-Pollak subindexes $P^m(p_m^0,p_m^1,u^0,x^0)$. A natural way of weighting these subindexes to form an overall cost-of-living index would be to use the base period expenditure share $p_m^0 \cdot x_m^0/p^0 \cdot x^0$ to weight the mth subindex $P^m(p_m^0,p_m^1,u^0,x^0)$. Call the resulting overall cost-of-living index P(0). Thus

$$\begin{aligned}
P(0) &\equiv \Sigma_{m=1}^M (p_m^0 \cdot x_m^0/p^0 \cdot x^0) P^m(p_m^0,p_m^1,u^0,x^0) \\
&= \Sigma_{m=1}^M (p_m^0 \cdot x_m^0/p^0 \cdot x^0)\, C^m(u^0,p_m^1,x^0)/C^m(u^0,p_m^0,x^0) \\
&= \Sigma_{m=1}^M (p_m^0 \cdot x_m^0/p^0 \cdot x^0)\, C^m(u^0,p_m^1,x^0)/p_m^0 \cdot x_m^0 \\
&\qquad\qquad \text{using (A6) in the appendix} \\
&\leq p^1 \cdot x^0/p^0 \cdot x^0 \equiv P_L \text{ using (A8) in the appendix.} \qquad (69)
\end{aligned}$$

Thus the aggregate two-stage index P(0) is bounded from above by the aggregate Laspeyres price index P_L. Note that if P^m were replaced by the Laspeyres group m subindex $p_m^1 \cdot x_m^0/p_m^0 \cdot x_m^0$, then (69) would collapse to P_L, the aggregate Laspeyres price index.

Consider now the Paasche-Pollak subindexes $P^m(p_m^0,p_m^1,u^1,x^1)$. A natural way of weighting these indexes to form an overall cost-of-living index would be to use period 1 expenditure shares $p_m^1 \cdot x_m^1/p^1 \cdot x^1$ as weights for the subindexes $P^m(p_m^0,p_m^1,u^1,x^1)$. However, instead of forming a simple weighted average, we shall form a harmonic mean:

$$P(1) \equiv [\Sigma_{m=1}^{M}(p_m^1 \cdot x_m^1/p^1 \cdot x^1)\ [P^m(p_m^0,p_m^1,u^1,x^1)]^{-1}]^{-1}$$

$$= [\Sigma_{m=1}^{M} p_m^1 \cdot x_m^1\ C^m(u^1,p_m^0,x^1)/p^1 \cdot x^1 C^m(u^1,p_m^1,x^1)]^{-1}$$

$$= [\Sigma_{m=1}^{M}\ C^m(u^1,p_m^0,x^1)/p^1 \cdot x^1]^{-1} \qquad \text{using (A7)}$$

$$\geq p^1 \cdot x^1/p^0 \cdot x^1 \equiv P_P \qquad \text{using (A9)}\,. \tag{70}$$

Thus the two stage-index P(1) is bounded from below by the aggregate Paasche index P_P. Note that if P^m were replaced by the Paasche group M subindex $p_m^1 \cdot x_m^1/p_m^0 \cdot x_m^1$, then (70) would collapse to P_P, the aggregate Paasche price index. Thus the overall Paasche price index P_P may be calculated as a weighted harmonic average of Paasche subindexes.

For each λ between 0 and 1, define the intermediate aggregate two-stage index $P(\lambda)$ by

$$P(\lambda) \equiv [\ \sum_{m=1}^{M}\ [(1-\lambda)s_m^0 + \lambda s_m^1]\ [P^m(p_m^0,p_m^1,(1-\lambda)u^0 + \lambda u^1,(1-\lambda)x^0 + \lambda x^1)]^r]^{1/r} \tag{71}$$

where $s_m^0 \equiv p_m^0 \cdot x_m^0/p^0 \cdot x^0$, $s_m^1 \equiv p_m^1 \cdot x_m^1/p^1 \cdot x^1$ and the exponent r that appears in (71) is defined by $r \equiv 1-2\lambda$.

It can be verified that P(0) and P(1) defined by (71) coincide with the P(0) and P(1) defined in (69) and (70) and moreover, $P(\lambda)$ is continuous for $0 \leq \lambda \leq 1$. Thus we may apply the usual Konüs [1924] proof (see the proof of Theorem 14) and conclude that there exists a λ^* between 0 and 1 such that

$$P_P \leq P(\lambda^*) \leq P_L \text{ or } P_L \leq P(\lambda^*) \leq P_P. \tag{72}$$

Thus $P(\lambda^*)$, a weighted average of the subindexes $P^m(p_m^0,p_m^1,(1-\lambda^*)\ u^0+\lambda^* u^1, (1-\lambda^*)x^0+\lambda^* x^1)$ for m=1,...,M has the usual overall Paasche and Laspeyres bounds and hence will be "close" to the one-stage Konüs cost-of-living index, $P_K(p^0,p^1,u^*)$, whose existence was given in Theorem 4 above.

Diewert [1978] showed that certain index number formulae, such as the superlative mechanical price index formulae $P_r(p^0,p^1,x^0,x^1)$ and $\tilde{P}_s(p^0,p^1,x^0,x^1)$ defined by (17) and (41), had a second order consistency in aggregation property.[24] Diewert's proof was highly technical and lacked economic motivation. Perhaps the results in this section cast some light on why "good" index number formulae may be expected to aggregate up to the "right" aggregate value when we use a "good" weighting scheme: the Paasche and Laspeyres bounds occur in an appropriate two-stage procedure as well as in the usual single-stage procedure for constructing a true cost-of-living index.

9. An Intertemporal Cost-of-Living Index

Pollak [1975a] developed the theory of subindexes and in Pollak [1975b], he attempted to apply his general theory of subindexes to the intertemporal context. However, in this section, we shall attempt to show that the general theory of subindexes cannot be very readily applied to the intertemporal context.

We begin with a single consumer who has a horizon that extends over T+1 Hicksian [1946] periods when we first observe his behaviour in period 0. Let his utility function be given by $F(x_0,x_1,\ldots,x_T)$ where x_t is the period t vector of purchases that the consumer plans to make in period t. Assume for simplicity that the consumer can borrow or lend a dollar from period 0 to period 1 at the interest rate r_1^0 and he **expects** the interest rate from period t to t+1 to be r_{t+1}^0 for t = 1, 2,..., T-1. Define the sequence of period 0 expected discount factors to be $\delta_1^0 \equiv 1/(1+r_1^0)$, $\delta_2^0 \equiv \delta_1^0/(1+r_2^0)$,..., $\delta_T^0 \equiv (\delta_{T-1}^0/1+r_T^0)$. Suppose that from the vantage point of period 0, the consumer **expects** the spot price vector $p_t^0 >> 0_N$ to prevail during period t for t = 1,2,...,T. The price vector $p_0^0 >> 0_N$ is the **observable** vector of market prices prevailing during period 0. Let p^0 be the vector of discounted expected prices

$$p^0 \equiv (p_0^0,\delta_1^0 p_1^0,\delta_2^0 p_2^0,\ldots,\delta_T^0 p_T^0).$$

We assume that $x^0 \equiv (x_0^0, x_1^0, \ldots, x_T^0)$ solves the period 0 expected utility maximization problem:

$$\max_x \left\{F(x)\colon p^0 \cdot x \le p^0 \cdot x^0 \equiv W^0\right\} \tag{73}$$

where $W^0 > 0$ is the consumer's initial wealth. Letting C be the cost function: we have the following equation (where $\delta_0^0 \equiv 1$):

$$\begin{aligned}
W^0 &= C[F(x^0), p^0] \\
&\equiv \min_x \left\{p^0 \cdot x\colon F(x) \ge F(x^0)\right\} \\
&= \min_{x_0,\ldots,x_T} \left\{ \sum_{t=0}^{T} \delta_t^0 p_t^0 \cdot x_t\colon F(x_0,\ldots,x_T) \ge F(x_0^0,\ldots,x_T^0)\right\} \\
&= \sum_{t=0}^{T} \delta_t^0 p_t^0 \cdot x_t^0 \\
&= \sum_{t=0}^{T} C^t[F(x^0), \delta_t^0 p_t^0, x^0] \qquad \text{using (A6)} \\
&= C^0[F(x^0), p_0^0, x^0] + \sum_{t=1}^{T} \delta_t^0 C^t[F(x^0), p_t^0, x^0] \qquad (74)
\end{aligned}$$

where (74) follows from the homogeneity properties of the conditional cost functions C^t which are defined in a manner analogous to C^m in (65). In period 0, we may observe the consumer's initial wealth W^0, the vector of prevailing market prices p_0^0, the consumer's vector of actual period 0 purchases x_0^0, his savings $W^0 - p_0^0 \cdot x_0^0$, and perhaps the *ex ante* period 0 interest rate r_1^0 (and the corresponding discount rate $\delta_1^0 \equiv 1/(1+r_1^0)$). *All of the other variables are unobservable.*

Now suppose that we can observe what happens during period 1. There will be new expectations about nominal interest rates r_t^1 for $t=2,\ldots,T$ and a new sequence of discount factors $\delta_t^1 \equiv \delta_{t-1}^1/(1+r_t^1)>0$, the consumer will have a new initial wealth W^1, there will be a new spot price vector for goods p_1^1 (which may or may not be equal to the consumer's period 0 expectation about this vector of spot prices p_1^0), and the consumer will have new expectations about future spot prices, $p_2^1,p_3^1,\ldots,p_T^1$. We now assume that $x_1^1,x_2^1,\ldots,x_T^1$ solves the following period 1 expected utility maximization problem:

$$\max_{x_1,\ldots,x_T} \left\{F(x_0^0,x_1,\ldots,x_T): p_1^1\cdot x_1^1 + \delta_2^1 p_2^1\cdot x_2 + \cdots + \delta_T^1\, p_T^1\cdot x_T \right. \tag{75}$$
$$\left. \leq p_1^1\cdot x_1^1 + \delta_2^1 p_2^1\cdot x_2^1 + \cdots + \delta_T^1\, p_T^1\cdot x_T^1 = W^1\right\}.$$

Note that the period 0 decision vector x_0^0 appears in (75), since we are stuck with it in period 1. Define $x^1 \equiv (x_0^0,x_1^1,\ldots,x_T^1)$. Proceeding in a manner analogous to our derivation of (74), we find that

$$W^1 \equiv C^1[F(x^1),p_1^1,x^1] + \sum_{t=2}^{T} \delta_t^1\, C^t[F(x^1),p_t^1,x^1]. \tag{76}$$

In period 1, we may observe the consumer's period 1 wealth W^1, the period 1 vector of actual market prices p_1^1, the consumer's purchases of goods during the period x_1^1, his saving $W^1 - p_1^1\cdot x_1^1$, and the *ex post* rate of return that was earned going from period 0 to period 1, r_1^1 (remember the corresponding *ex ante* expected rate of return was r_1^0).

Using the techniques outlined in the previous section, we may establish the following observable bounds for two intertemporal subindexes:

$$\min_i \left\{p_{1i}^1/p_{0i}^0\right\} \leq \frac{C^0[F(x^0),\, p_1^1,x^0]}{C^0[F(x^0),p_0^0,x^0]} \leq p_1^1\cdot x_0^0/p_0^0\cdot x_0^0 \equiv P_L;$$

$$P_P \equiv p_1^1\cdot x_1^1/p_0^0\cdot x_1^1 \leq \frac{C^1[F(x^1),\, p_1^1,\, x^1]}{C^1[F(x^1),\, p_0^0,\, x^1]} \leq \max_i \left\{p_{1i}^1/p_{0i}^0\right\}.$$

However, the bounds established above do not allow us to answer any really interesting questions.

Although we cannot compute a complete intertemporal cost-of-living index in general, it is possible to make a comparison of the expected welfare of the consumer during periods 0 and 1. Recall that the *ex post* period 0 rate of return was defined to be r_1^1. Define the *ex post* discount rate $\delta_1^1 \equiv 1/(1+r_1^1)$. Then we may rewrite the period 1 expected utility maximization problem (75) in the following manner:

$$\max_{x_1,\ldots,x_T} \{F(x_0^0,x_1,\ldots,x_T): p_0^0 \cdot x_0^0 + \delta_1^1 p_1^1 \cdot x_1 + \delta_1^1 \delta_2^1 p_2^1 \cdot x_2 + \ldots \tag{77}$$

$$+ \delta_1^1 \delta_T^1 p_T^1 \cdot x_T \le p_0^0 \cdot x_0^0 + \delta_1^1 w^1\}.$$

The objective functions in (75) and (77) are the same and the constraint in (77) is obtained by multiplying both sides of the constraint in (75) by δ_1^1 and then adding the constant $p_0^0 \cdot x_0^0$ to both sides.

The consumer's period 0 expected utility maximization problem (73) may be rewritten as (78) when we fix $x^0 = x_0^0$:

$$\max_{x_1,\ldots,x_T} \{F(x_0^0,x_1,\ldots,x_T): p_0^0 \cdot x_0^0 + \delta_1^0 p_1^0 \cdot x_1 + \delta_2^0 p_2^0 \cdot x_2 \tag{78}$$

$$+ \ldots + \delta_T^0 p_T^0 \cdot x_T \le w^0\}.$$

If the consumer's period 0 expectations equal his period 1 expectations about future prices and interest rates and if the expected *ex ante* first period rate of return r_1^0 equals the *ex post* rate of return r_1^1, then it can be verified that the prices appearing in the constraint of (77) are identical to the prices appearing in the constraint of (78). Under these conditions, we could say that the choice set of the consumer (and hence his welfare) has increased going from period 0 to period 1 if

$$p_0^0 \cdot x_0^0 + \delta_1^1 w^1 > w^0 \tag{79}$$

where $\delta_1^1 \equiv 1/(1+r_1^1)$ and r_1^1 is the *ex post* one-period rate of return on assets going from period 0 to period 1.

The criterion for an expected welfare increase (or for an increase in real wealth) may be rewritten as

$$W^1 > (W^0 - p_0^0 \cdot x_0^0)(1 + r_1^1)$$

$$= \text{(first period savings)}(1 + \text{ex post rate of return}).$$

This criterion for an increase in real wealth is due to Hicks [1946; p.175]. However, its validity does require the assumption of constant expectations, an assumption that is unlikely to be fulfilled under present economic conditions.

The problem of measuring welfare changes in a general Hicksian intertemporal choice model seems to be inherently difficult. Virtually all of the index number techniques that we have surveyed and developed in this paper rely on the twin assumptions of optimizing behaviour on the part of the consumer and observability of market prices and the consumer's quantity choices (or we require assumptions about the constancy of expectations that are unlikely to be met in practice). Hence in order to apply the traditional theory of index numbers in the intertemporal context, it appears to be necessary to place *a priori* restrictive assumptions on the form of the intertemporal utility function $F(x_0,x_1,\ldots,x_T)$ such as intertemporal additivity;[25] i.e., $F(x_0,x_1,\ldots,x_T) = \Sigma_{t=0}^{T} f(x_t)$. If we do this, then we may apply traditional index number theory in order to measure changes in the one-period utility function $f(x_t)$; i.e., we could approximate the change in $f(x_1^1)/f(x_0^0)$ using the period 0 and 1 price and quantity data in the usual manner.

10. Spatial Cost-of-Living Indexes

The basic problem to be considered in this section is the problem of comparing the level of prices in different cities or localities. The problem is isomorphic to the usual problem of making international comparisons.[26]

If we are willing to assume that a certain class of consumers in one location (location 0 say) has the same one-period preferences as another class of consumers in another location (location 1 say) and if there are no significant differences in environmental (non-market) variables in the two locations, then we may simply apply the theory outlined in Sections 2-4 above, where the superscripts 0 and 1 will now refer to locations. However, there were two rather big "ifs" in the previous sentence.

We may relax the restrictiveness of the second "if" by using a Pollak subindex (recall Section 8) of the form

$$C^1[F(x^0),p_1^1,x_2^0]/C^1[F(x^0),p_1^0,x_2^0] \equiv P^1(p_1^0,p_1^1,F(x^0),x_2^0) \text{ and} \tag{80}$$

$$C^1[F(x^1),p_1^1,x_2^1]/C^1[F(x^1),p_1^0,x_2^1] \equiv P^1(p_1^0,p_1^1,F(x^1),x_2^1] \tag{81}$$

where P^1 is a Pollak subindex of the cost of living over market goods (recall (66)), p_1^0 and p_1^1 are vectors of observed market prices in locations 0 and 1 respectively, $F(x) = F(x_1,x_1)$ is the consumer's utility function defined over combinations of market goods x_1 and non-market locational amenities x_2, x_1^0 and x_1^1 are the observed market choice vectors for consumers 0 and 1 respectively, x_2^0 is the amenity vector in location 0, x_2^1 is the amenity vector in location 1, $x^0 \equiv (x_1^0,x_2^0)$ and $x^1 \equiv (x_1^1,x_2^1)$.

The bounds derived in Section 8 (see (67) and (68)) are applicable to the subindexes defined in (80) and (81) under the usual optimizing behaviour assumptions:

$$\min_i \{p_{1i}^1/p_{1i}^0\} \leq P^1(p_1^0,p_1^1,F(x^0),x_2^0) \leq p_1^1 \cdot x_1^0/p_1^0 \cdot x_1^0 \equiv P_L \text{ and} \tag{82}$$

$$P_P \equiv p_1^1 \cdot x_1^1/p_1^0 \cdot x_1^1 \leq P^1(p_1^0,p_1^1,F(x^1),x_2^1) \leq \max_i \{p_{1i}^1/p_{1i}^0\}. \tag{83}$$

Theorem 19 may be applied in the present context as well.

The assumption that the consumers in the two locations have the same preferences may also be relaxed; see Caves, Christensen and Diewert [1982b; p.1410] and Denny and Fuss [1983] for various "translog approaches".

11. Leisure and Labour Supply in the Cost-of-Living Index

Consider a household that can supply various kinds of labour service. It is natural to assume that the household has one-period preferences defined over various combinations of market goods $x \equiv (x_1,\ldots,x_N) \geq 0_N$ and labour supplies $y \equiv (y_1,\ldots,y_M) \leq 0_M$ where $y_m \leq 0$ is the **negative** of the number of hours of the mth type of work supplied by the household. The preferences are summarized by the utility function $F(x,y)$.[27] Suppose now that the household faces the positive commodity price vector $p >> 0_N$ and the positive (after tax marginal) wage vector $w >> 0_M$. Then we may define the household's cost function C in the usual manner:

$$C(u,p,w) \equiv \min_{x,y} \left\{ p \cdot x + w \cdot y : F(x,y) \geq u \right\}. \tag{84}$$

Given period 0 and 1 price vectors, (p^0,w^0) and (p^1,w^1), we may be tempted to define the Konüs cost-of-living index in the usual manner:

$$P_K(p^0,w^0,p^1,w^1,u) \equiv C(u,p^1,w^1)/C(u,p^0,w^0). \tag{85}$$

If we could be assured that $C(u,p,w) > 0$ for $u>0$, and $p >> 0_N$, and $w >> 0_M$, then there would be no problem with definition (85), and in fact we could derive the usual bounds that we derived in Section 2. However, $C[F(x^0,y^0), p^0,w^0] > 0$ corresponds to a situation where the value of household consumption in period 0, $p^0 \cdot x^0$ exceeds the value of household labour supply, $-w^0 \cdot y^0$. There is no reason for this to be the case. Worse yet, the two values could coincide (i.e., we could have $p^0 \cdot x^0 + w^0 \cdot y^0 = 0$) in which case $P_K(p^0,w^0,p^1,w^1, F(x^1,y^1))$ becomes undefined (since we are dividing by zero in definition (85)).

We could attempt to avoid these problems by assuming that the household has preferences defined over different combinations of market goods x and leisure, where the leisure vector is defined by $b + y \geq 0_M$ and $y \leq 0_M$ represents hours of work and $b >> 0_M$ is a positive vector that could perhaps represent the maximum hours that could be supplied of the various types of household labour services. This approach is pursued in Riddell [1983]. However, it will be difficult to come to an agreement on just what value we should take

for the vector b (particularly if members of the household are holding multiple jobs). Hence the resulting Konüs Cost of Living index would be somewhat arbitrary, and should therefore not be used as an inflation measure.[28].

However, it is still possible to use index number techniques in order to obtain approximations to the change in the household's welfare which occurred going from period 0 to period 1. Essentially, what we shall do is measure welfare changes in terms of proportional changes in consumption goods.

First we define the household's conditional consumption deflation function D by

$$D(u,x,y) \equiv \max_k \left\{k: F(x/k,y) \geq u,\ k > 0\right\} \qquad (86)$$

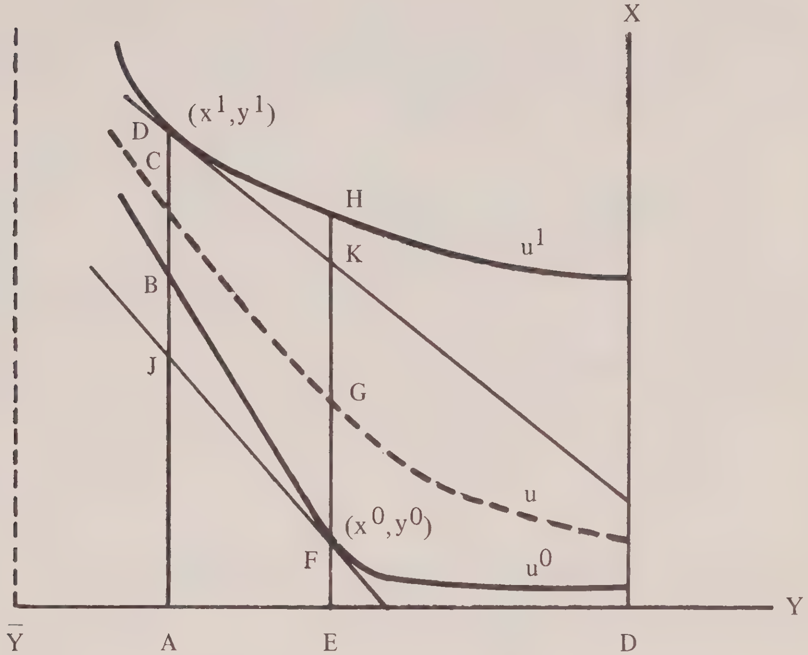
where $u > 0$ is a reference utility level, $x > 0_N$ is a consumption vector, $y \leq 0_M$ is a vector of labour supplies indexed negatively and F is the household utility function. If $D(u,x,y,)$ > 1 (<1), then the household joint consumption labour supply vector yields a utility level greater (less) than u, while if $D(u,x,y) = 1$, then (x,y) yields precisely the utility level u. In general $D(u,x,y)$ tells us the proportion $k^* > 0$ that we have to deflate the consumption vector x so that $(x/k^*,y)$ will yield utility level u for the household.

A Malmquist [1953] consumption quantity index may now be defined in the usual manner:

$$Q(x^0,y^0,x^1,y^1,u) \equiv D(u,x^1,y^1)/D(u,x^0,y^0). \qquad (87)$$

A geometric interpretation of the quantity index defined by (87) may be found on Figure 3 for the case of one consumption good ($N = 1$) and one type of labour supply ($M = 1$).

Figure 3 The Malmquist Quantity Index in the Labour Supply Context.



For the reference utility level u (which is between $u^0 \equiv F(x^0,y^0)$ and $u^1 \equiv F(x^1,y^1)$), it can be verified that $Q(x^0,y^0,x^1,y^1,u) = [DA/CA]/[FE/GE]$.

As usual, it is useful to let the references utility level u in (87) be either the base utility level $u^0 \equiv F(x^0,y^0)$ or the period 1 utility level $u^1 \equiv F(x^1,y^1)$. Thus define the *Laspeyres-Malmquist consumption quantity index* by $Q(x^0,y^0,x^1,y^1,F(x^0,y^0))$ and the *Paasche-Malmquist consumption quantity index* by $Q(x^0,y^0,x^1,y^1,F(x^1,y^1))$. In Figure 2, the first of these indexes is DA/BA while the second is HE/FE.

Many of the theorems that were stated in Section 3 go through in the present context. In particular, a counterpart to Theorem 7 is true, so if the reference utility level in (87) is between u^0 and u^1, $Q(x^0,y^0,x^1,y^1,u) > 1$ $(<1, =1)$ indicates that household utility has increased (decreased, remained constant) going from period 0 to period 1.

We also obtain the following counterpart to Theorem 9.

Theorem 21

Suppose F satisfies the modified Conditions I and the household's observed period 0 and period 1 choices (x^0,y^0) and (x^1,y^1) are consistent with utility maximizing or cost minimizing behaviour; i.e., x^0,y^0 satisfies

$$p^0 \cdot x^0 + w^0 \cdot y^0 = C[F(x^0,y^0),p^0,w^0] \tag{88}$$

where (p^0,w^0) are the observed positive period 0 prices and C is the cost function defined by (84), and (x^1,y^1) satisfies

$$p^1 \cdot x^1 + w^1 \cdot y^1 = C\,[F(x^1,y^1),\, p^1,w^1] \tag{89}$$

where (p^1,w^1) are the observed positive period 1 prices. Then provided that $p^0 \cdot x^0 + w^0 \cdot (y^0 - y^1) > 0$ and $p^1 \cdot x^1 + w^1 \cdot (y^1 - y^0) > 0$,

$$Q(x^0,y^0,x^1,y^1,\, F(x^0,y^0)) \leq p^0 \cdot x^1/[p^0 \cdot x^0 + w^0 \cdot (y^0 - y^1)] \equiv \alpha \text{ and} \tag{90}$$

$$Q(x^0,y^0,x^1,y^1,\, F(x^1,y^1)) \geq [p^1 \cdot x^1 + w^1 \cdot (y^1 - y^0)]/p^1 \cdot x^0 \equiv \beta. \tag{91}$$

The bounds in (90) and (91) may be illustrated by referring to Figure 3. (90) becomes DA/BA ≤ DA/JA and (91) becomes HE/FE ≥ KE/FE. As the reader can observe, the bounds are rather close to the appropriate theoretical index. It is also true that the bounds α and β will often be close to each other. Hence the following counterpart to Theorem 10 is of some practical interest.

Theorem 22

Assume the regularity conditions of the previous theorem and define the base utility $u^0 \equiv F(x^0,y^0)$ and the period 1 utility level $u^1 \equiv F(x^1,y^1)$. Then there exists a reference utility level u^* between u^0 and u^1 such that the Malmquist consumption quantity index $Q(x^0,y^0,x^1,y^1,u^*)$ lies between α and β, the bounds defined in (90) and (91).

Thus if α and β are close, we will be able to obtain a good estimate of $Q(x^0,y^0,x^1,y^1,u^*)$ by averaging α and β.

For an empirical implementation of the above material, see Riddell [1983].

12. Durables in the Cost-of-Living Index

The treatment of consumer durables in the CPI and in the true cost-of-living index is an interesting and controversial issue.[29]

The basic issues can readily be explained. Consider a durable good that can be purchased at the beginning of the current period at the spot price q^0. A consumer can purchase this good at the beginning of the period, use the good during period 0, but because of the good's durable nature, some of it will be left over at the beginning of the following period. The consumer could sell his used durable good at a (possibly hypothetical) second-hand market at an **expected** price of q^{01}. Assuming that the consumer can lend or borrow at the rate of return r, we may follow Hicks [1946] and compute the present value of the cost of buying one unit of the good, using it for one period, and selling it next period. The resulting **user cost** p is

$$p = q^0 - q^{01}/(1+r) = [q^0 r + (q^0 - q^{01})]/[1+r]. \qquad (92)$$

The first term on the right-hand side of (82) is an interest cost while the second term combines the effects of anticipated capital gains and depreciation. We can separate out these two effects if we let q^1 be the spot price of a new unit of the durable that the consumer expects to prevail during period 1. Hence we may write $(q^0\text{-}q^{01}) = (q^1\text{-}q^{01})\text{-}(q^1\text{-}q^0)$ = depreciation - capital gains. Thus the user cost (92) becomes

$$p = (1+r)^{-1}[q^0 r + (q^1 - q^{01}) - (q^1 - q^0)]. \qquad (93)$$

The above derivation of the user cost of a durable in discrete time was essentially obtained by Diewert [1974; p.504] and Pollak [1975b]. The term $(1+r)^{-1}$ may strike the reader

as being a bit odd, but we need it so that when $q^{01}=0$ (so that the good is actually a non-durable), then the user cost p collapses to the period 0 purchase price q^0.

Unfortunately, there are many problems associated with the user cost formulae (92) or (93).

The first problem is that the prices q^{01} and q^1 are not market prices - they are the consumer's period 0 expectations of what the market prices will actually be in period 1. It is not realistic to assume that our non-renting consumer will be able to accurately forecast these future prices. If consumers were able to accurately forecast future prices, then the existing rental market price for the services of the durable would equal the user cost p defined by (92) where the expected price q^{01} in (92) is replaced by the observed *ex-post* market price. Hendershott [1980; 406] demonstrates that rental prices for houses do not track ex post user costs for housing very closely.[30] However, all is not lost if there are some rental markets for the class of durable goods under consideration.

If we do have some rental market information on the durable then we could assume that the rental price equals the user cost p that appears in (92) and we could use (92) to **solve** for the expected price q^{01} in terms of the observed market prices p,q^0 and r. Thus expectational information gleamed in this way for classes of durables that have rental markets could be applied to forecast expectations of future prices for classes of "similar" durables that do not have rental markets.

Most goods can only be purchased in integral numbers, and for most goods, this does not cause major problems. However, some durable goods such as cars and houses may be purchased only in integer units, and such purchases would form a large share of the consumer's total expenditure. Hence we cannot neglect the lumpiness problem for such classes of durables. How may we apply traditional "continued" utility and index number theory to this situation?

A possible solution is illustrated in Figure 4 below. Assume x_2 represents units of a "continuous" good, ice cream say, while x_2 represents the number of television sets that a household holds during a period.

In Figure 4, we have graphed two indifference "curves" for the household. Only the "kinky" points on the curves plus the line segments parallel to the x_2 axis are actual feasible choices for the household, but for all practical purposes, we can replace the original preferences defined only over integer combinations of TV sets by continuous preferences with "kinks".[31] The resulting preference function F(x) may be treated in the normal manner as far as index number theory is concerned. Note that the economic effect of the "kinks" will be to make the consumer change his durable holdings only after relatively large changes in the rental prices of the durables relative to non-durable goods; i.e., responses will be "sticky". This point should be taken into account in econometric work, but it need not concern us from the viewpoint of index number theory.

Another difficulty (which **does** create problems for us from the viewpoint of index number theory) is that the expected buying and selling price of the durable may not be the same; i.e., there may be significant transactions costs associated with buying and selling units of the durable. The effect of this difference in buying and selling prices for the durable will be to put a "link" in the consumer's budget constraint around his initial holdings of the durable. See Figure 5 below.

Figure 4.

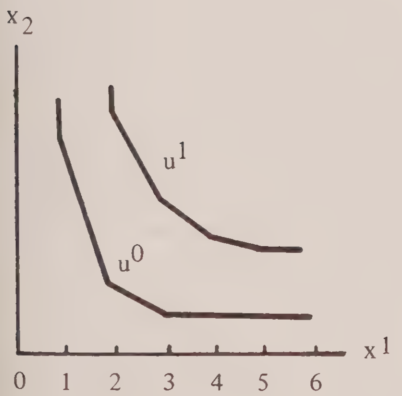
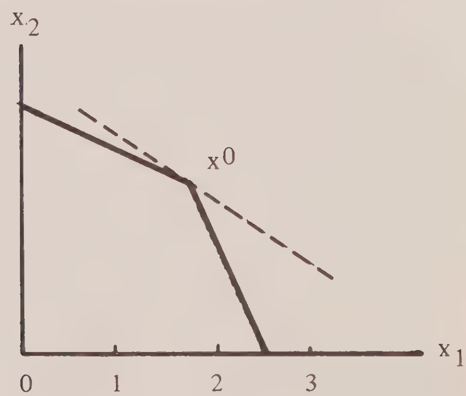


Figure 5.



If the consumer is observed at x^0 during period 0, the correct price of the durable to use in an index number comparison lies somewhere between the buying and selling price. For our purposes, we shall probably have to settle for taking an average of the two prices. Note that the effect of this difference in buying and selling prices will again have the effect of making the consumer's demand to hold durables ''sticky'' around his initial holdings.[32]

Differences in borrowing and lending rates for a household may also have the effect of introducing a ''kink'' into the budget set when the purchase of a large consumer durable is contemplated. Deaton and Muellbauer [1980; Ch.13] discuss the effect of a borrowing constraint, and in fact, they have an excellent discussion of the problems involved in modelling the demand for consumer durables.

The final problem that we wish to discuss is the problem of calculating the user cost of a durable when there are tax considerations involved.

At the marginal investment point, we assume that the durable holder invests in a market asset that earns a before tax rate of return r. Suppose that consumer's marginal tax rate is τ. The present value of the cost of using the durable for one period is

$$p \equiv q^0 - \frac{q^{01}}{1+(1-\tau)r} + \frac{\tau\,\alpha(q^{01}-q^0)}{1+(1-\tau)r} - \frac{\tau\beta r_M q^0}{1+(1-\tau)r} + \frac{tq^0}{1+(1-\tau)r} - \frac{\tau\delta q^0}{1+(1-\tau)r} \tag{94}$$

where q^0 and q^{01} are the same as before, α is the proportion of capital gains on the durable that is taxable, β is the proportion of (mortgage) interest that is deductible from taxable income (and r_M is the appropriate mortgage interest rate), t is a user tax rate on holdings of the durable (e.g., a property tax), δ is the depreciation rate that is allowed for taxation purposes, and all taxes are assumed to be payable in the following period (and hence they are discounted). If we assume $\alpha=\beta=\delta=0$, then (94) reduces to

$$p = [(1-\tau)rq^0 + tq^0 - (q^{01}-q^0)]/[1+(1-\tau)r]. \tag{95}$$

Thus the higher is the marginal tax rate , the lower will be the user cost p. If the durable holder is a borrower (at the mortgage rate r_M say), then the discount factor in (94) and (95) must be replaced by $1 + r_M$. Under these conditions, (95) becomes

$$p = [r_M q^0 + tq^0 - (q^{01} - q^0)]/[1 + r_M]. \tag{96}$$

which will be much higher than the user cost defined by (95).

Thus the tax and financial situation of the consumer plays an essential role in the calculation of user costs for durables. This is a very unfortunate result, since the informational requirements for implementing a user cost formula such as (94) are very high. In particular, the expected price q^{01} that appears in (94) is not directly observable. Also it may be very difficult to determine precisely what is the appropriate opportunity cost of the marginal investment (r) for the consumer under consideration. These problems are particularly acute in the case of housing, since housing expenditures are generally a large proportion of a typical household budget. In order to avoid the difficult measurement problems associated with the user cost approach, Gillingham [1982] suggests essentially that the price quantity data that pertains to the rental segment of the housing market be extrapolated to the entire housing market. The problem with this rather sensible suggestion is that the rental segment of the housing market is generally not representative of the entire housing market. A possible reason for this non-representativeness emerges if we compare the user cost for a house for a rich individual with a high marginal tax rate δ (see (95)) with the user cost formula for a less well-off individual who has no non-labour income and holds a mortgage on his house (recall (96)): the same house will cost the rich individual far less in terms of user cost than the poor individual. Furthermore, the tax laws in most Western countries will generally make it more profitable for a rich individual to own and live in his house rather than rent out his house and live in a rental house of comparable quality. Thus the rental price information for the rental portion of the housing market will not be representative for the market as a whole (due to these tax considerations). Moreover, as nominal interest rates and marginal tax rates changed over time, we could expect the price of rental housing to systematically deviate from the appropriate user cost index of non-rental housing. Thus I do not believe that the informational difficulties that are inherent in the user

cost formula (94) can be avoided. For further discussion on the role of tax considerations in the construction of user costs, see Darrough [1983].

We conclude this section with a reminder that all of our bounds on the consumer's true cost-of-living index rested on the assumption of utility maximizing behaviour subject to a budget constraint. The prices that appear in the budget constraint are observable market prices in the case of non-durable goods, but for a durable good which is owned by the consumer, the appropriate price must be an *ex ante* user cost of the form (94), which depends on the (unobservable) anticipated price of the depreciated durable which is expected to prevail in the following period. It is not in general appropriate to use an *ex post* user cost formula of the form (94) where the anticipated price q^{01} is replaced by an (observable) market price, since the resulting *ex post* user cost may well be negative if there is an unanticipated inflation in the price of the durable. Thus the existence of unanticipated inflation (or deflation) and the non-neutrality of the income tax with respect to the treatment of durables make it very difficult to construct a Konüs cost-of-living index (or bounds to it) for a household. Thus there are costs due to unanticipated inflation and the non-neutrality of the present tax system.

13. The New Goods Problem

The standard approach to the new goods problem in the context of consumer theory dates back to Hicks [1940]: the period after a new good appears (period 1 say), we attempt to impute a price for the new good in period 0 which would just make the consumer's demand for the good equal to 0 in period 0. The details of an econometrically implementible approach are outlined in Diewert [1980; pp.501-503].

However, a simpler method that is more practical is readily available: as soon as a new good appears, start collecting price and quantity information on it. Since initially the quantity purchased will be low and the price will often be rather high, a quick introduction of the good into the CPI universe of prices would solve most of the practical problems associated with the current neglect of the quality change problem in official CPIs.[33] It is true that many new goods quickly disappear, but this causes no particular theoretical problems. However, we must concede that linking the prices of similar new goods that vary

in quality poses some practical problems.

14. Conclusion

The main conclusions which emerge from this paper are listed below.

(i) In addition to the usual Laspeyres based CPI, a Paasche based CPI should also be published as frequently as possible, since an appropriate true cost-of-living index lies between a Paasche and Laspeyres index. This means that household surveys where quantity information is collected should be undertaken more frequently, say every second year.

(ii) CPIs should be constructed on a more disaggregated basis (by household demographic characteristics, by income and by region).

(iii) Labour supply and leisure should be introduced into the CPI framework on an experimental basis. The consumer-worker's income tax position will play an important role here.

(iv) Intertemporal CPIs are too problematical to be introduced at this time.

(v) The treatment of consumer durables, particularly housing, is not very satisfactory at present. Various user cost and rental equivalent alternatives should be tried on an experimental basis. The importance of the household's tax and financial position should not be overlooked.

(vi) The treatment of seasonal commodities is also unsatisfactory. A more satisfactory treatment of seasonal commodities from the viewpoint of economic theory is outlined in Diewert [1983b].

(vii) New goods should be introduced into the CPI universe of prices as soon as possible. The neglect of new goods provides an upward bias to the existing CPI of a possibly major magnitude.

If the above recommendations are implemented, then not only will we have much more accurate information on how inflation affects different consumer groups, we will also be able to simulate how changes in tax policy affect the welfare of the different household groups.

Appendix: Proofs of Theorems

Proof of Theorem 1: Define $C(u,p) \equiv uP(p^*, p,u)$ for $u>0$ and $p>>0_N$. Let $u>0$, $p^0>>0_N$ and $p^1>>0_N$. Then

$$C(u,p^1)/C(u,p^0) \equiv uP(p^*,p^1,u)/uP(p^*,p^0,u)$$

$$= P(p^*,p^1,u)/P(p^*,p^0,u)$$

$$= P(p^0,p^*,u)P(p^*,p^1,u)$$

using the time reversal property (ii)

$$= P(p^0,p^1,u)$$

using the circularity property (iii).

Note that properties (ii) and (iii) imply $P(p^*,p^*,u) = 1$. Hence $C(u,p^*) \equiv uP(p^*,p^*,u) = u$ for all $u > 0$ which is (3).

The proof of the converse part of the theorem is straightforward. I owe this method of proof to David Donaldson.

Proof of Theorem 14: Define $h(\lambda) \equiv P_D(p^0,p^1, (1-\lambda)u^0+\lambda u^1)$ for $0 \le \lambda \le 1$. Note that $h(0) = P_D(p^0,p^1,u^0)$ and $h(1) = P_D(p^0,p^1,u^1)$. There are 24 possible *a priori* inequality relations that are possible between the four numbers $h(0)$, $h(1)$, $\bar{P}_L$ and $\bar{P}_P$. However, (51) and (52) imply that $h(0) \le \bar{P}_L$ and $\bar{P}_P \le h(1)$. This means that there are only six

possible inequalities between the four numbers:

(1) $h(0) \leq \bar{P}_L \leq \bar{P}_P \leq h(1)$, (2) $h(0) \leq \bar{P}_P \leq \bar{P}_L \leq h(1)$, (3) $h(0) \leq \bar{P}_P \leq h(1) \leq \bar{P}_L$, (4) $\bar{P}_P \leq h(0) \leq \bar{P}_L \leq h(1)$, (5) $\bar{P}_P \leq h(1) \leq h(0) \leq \bar{P}_L$ and (6) $\bar{P}_P \leq h(0) \leq h(1) \leq \bar{P}_L$. Since the individual cost functions $C^h(u_h,p^t)$ are continuous in u_h, it can be seen that $P_D(p^0,p^1,u)$ defined by (49) is continuous in the vector of utility variables u. Hence $h(\lambda)$ is a continuous function for $0 \leq \lambda \leq 1$ and assumes all intermediate values between h(0) and h(1). By inspecting cases (1) to (6) above, it can be seen that we can choose λ between 0 and 1 (call this number λ^*) so that $\bar{P}_L \leq h(\lambda^*) \leq \bar{P}_P$ for case (1) or so that $\bar{P}_P \leq h(\lambda^*) \leq \bar{P}_L$ for cases (2) to (6). Now define $u^* \equiv (1-\lambda^*)u^0 + \lambda^* u^1$ and the proof is complete.

Proof of Theorem 15: Using expression (54) for $P_{PP}(p^0,p^1,u^0)$ and $P_{PP}(p^0,p^1,u^1)$, noting that the shares $s^h(u^0,p^0)$ and $s^h(u^1,p^0)$ defined by (55) are non-negative and sum to 1, and using Theorem 2, we may readily establish the left inequality in (57) and the right inequality in (58).

Consider now the right inequality in (57). This follows readily from definition (53), the equalities $C^h(u_h^0,p^0) = p^0 \cdot x_h^0$ for $h=1,\ldots,H$ and the inequalities

$$\begin{aligned} C^h(u_h^0,p^1) &\equiv \min_x \{p^1 \cdot x: F^h(x) \geq F^h(x_h^0) \equiv u_h^0\} \\ &\leq p^1 \cdot x_h^0 \end{aligned} \tag{A1}$$

which follow since x_h^0 is feasible for the minimization problem.

The left inequality in (58) follows from definition (53) when $u=u^1$, the equalities $C^h(u_h^1,p^1)=p^1 \cdot x_h^1$, and the inequalities

$$\begin{aligned} C^h(u_h^1,p^0) &\equiv \min_x \{p^0 \cdot x: F^h(x) \geq F^h(x_h^1) \equiv u_h^1\} \\ &\leq p^0 \cdot x_h^1. \end{aligned} \tag{A2}$$

Proof of Theorem 16: The proof of this theorem is identical to the proof of Theorem 14 if we replace P_D by P_{PP}, $\bar{P}_L$ by P_L and $\bar{P}_P$ by P_P.

Proof of Theorem 17: First note that for each household h,

$$
\begin{aligned}
D^h(u_h^0, x_h^1) &\equiv \max_k \{k: F^h(x_h^1/k) \geq u_h^0\} \\
&= k_h^1 \text{ where } F^h(x_h^1/k_h^1) = u_h^0.
\end{aligned}
$$

Thus

$$
\begin{aligned}
p^0 \cdot x_h^0 &= C^h(u_h^0, p^0) \qquad \text{(A3)} \\
&= \min_x \{p^0 \cdot x: F^h(x) \geq u_h^0\} \\
&\leq p^0 \cdot x_h^1/k_h^1
\end{aligned}
$$

since x_h^1/k_h^1 is feasible for the cost minimization problem. Thus (A3) yields the following inequality:

$$D(u_h^0, x_h^1) \leq p^0 \cdot x_h^1 / p^0 \cdot x_h^0 \qquad \text{for } h = 1, \ldots, H.$$

Repeating the above argument interchanging the superscripts 0 and 1 yields the following inequality:

$$D(u_h^1, x_h^0) \leq p^1 \cdot x_h^0 / p^1 \cdot x_h^1 \qquad \text{for } h = 1, \ldots, H. \qquad \text{(A5)}$$

From definition (60), we have

$$
\begin{aligned}
W(X^0,X^1,u^0) &\equiv \Sigma_{h=1}^H \beta_h^0 D^h(u_h^0,x_h^1)/D^h(u_h^0,x_h^0) \\
&= \Sigma_{h=1}^H \beta_h^0\, D^h(u_h^0,x_h^1) \qquad \text{since } D^h(u_h^0,x_h^0) = 1 \\
&\le \Sigma_{h=1}^H \beta_h^0\, p^0 \cdot x_h^1/p^0 \cdot x_h^0
\end{aligned}
$$

using (A4) and $\beta_h^0 \ge 0$, which establishes (63). Similarly

$$
\begin{aligned}
W(X^0,X^1,u^0) &\equiv \Sigma_{h=1}^H \beta_h^1 D^h(u_h^1,x_h^1)/D^h(u_h^1,x_h^0) \\
&= \Sigma_{h=1}^H \beta_h^1\, 1/D^h(u_h^1,x_h^0) \qquad \text{since } D^h(u_h^1,x_h^1) = 1 \\
&\ge \Sigma_{h=1}^H \beta_h^1\, p^1 \cdot x_h^1/p^1 \cdot x_h^0 \qquad \text{using (A5) and } \beta_h^1 \ge 0.
\end{aligned}
$$

Proof of Theorem 18: Define $h(\lambda) \equiv W(X^0,X^1, (1-\lambda)u^0 + \lambda u^1)$ and repeat the proof of Theorem 14, where $\Sigma_{h=1}^H \beta_h^0 Q_L^h$ replaces $\bar{P}_L$ and $\Sigma_{h=1}^H \beta_h^1 Q_P^h$ replaces $\bar{P}_P$.

Proof of Theorem 19: First note that

$$
\begin{aligned}
C(F(x^0),p^0) &\equiv \min_x \left\{p^0 \cdot x : F(x) \ge F(x^0)\right\} \\
&= p^0 \cdot x^0 \qquad \text{by (5)} \\
&= \Sigma_{m=1}^M p_m^0 \cdot x_m^0 \\
&= \min_{x_1} \left\{p_1^0 \cdot x_1 + \sum_{m=2}^{M} p_m^0 \cdot x_m^0 : F(x_1,x_2^0,\ldots,x_M^0) \right. \\
&\qquad\qquad \left. \ge F(x_1^0,x_2^0,\ldots,x_M^0)\right\} \\
&\equiv C^1(F(x^0),p_1^0,x^0) + \Sigma_{m=2}^M p_m^0 \cdot x_m^0 \,.
\end{aligned}
$$

Hence $C^1(F(x^0), p_1^0, x^0) = p_1^0 \cdot x_1^0$. In a similar manner, we obtain the following equalities:

$$C^m(F(x^0), p_m^0, x^0) = p_m^0 \cdot x_m^0, \qquad m = 1,2,\ldots,M \text{ and} \tag{A6}$$

$$C^m(F(x^1), p_m^1, x^1) = p_m^1 \cdot x_m^1, \qquad m = 1,2,\ldots,M. \tag{A7}$$

Since x_m^0 is feasible for the conditional cost minimization problem defined by $C^m(F(x^0), p_m^1, x^0)$, we have

$$C^m(F(x^0), p_m^1, x^0) \le p_m^1 \cdot x_m^0, \qquad m = 1,\ldots,M. \tag{A8}$$

Since x_m^1 is feasible for $C^m(F(x^1), p_m^0, x^1)$,

$$C^m(F(x^1), p_m^0, x^1) \le p_m^0 \cdot x_m^1, \qquad m = 1,\ldots,M. \tag{A9}$$

The definition of $P^m(p_m^0, p^1, u^0, x^0) \equiv C^m(F(x^0), p_m^1, x^0/C^m(F(x^0), p_m^0, x^0)$, (A6) and (A8) yield (67) while $P^m(p_m^0, p_m^1, u^1, x^1) \equiv C^m(F(x^1), p_m^1, x^1)/C^m(F(x^1), p_m^0, x^1)$, (A7) and (A9) yield (68).

Footnotes

[1] The conceptual framework for the Canadian CPI is nicely explained in Statistics Canada [1982]. The CPI and the Implicit Consumption Price Index for Canada are compared and contrasted in Loyns [1972], who was mainly interested in their inflation measuring capabilities. My focus will be more welfare-oriented.

[2] Unfortunately, much of the material presented in this paper is a bit technical. Two useful references that lead the reader into the technical aspects of index number and growth measurement theory in a gentle fashion are Allen [1975], and Usher [1980]. More technical discussions may be found in Konüs [1924], Samuelson [1947; pp.146-162], Malmquist [1953], Pollak [1971], Afriat [1977] and Diewert [1981].

[3] Notation: x^T denotes the transpose of the column vector x, $p^Tx = p \cdot x \equiv \Sigma_{n=1}^N p_n x_n$ denotes the inner product of the vectors p and x, $x \geq 0_N$ means each component of the vector x is non-negative, $x >> 0_N$ means each component if positive, and $x > 0_N$ means $x \geq 0_N$ but $x \neq 0_N$.

[4] F is a function defined over the non-negative orthant $\{x: x \geq 0_N\}$ that has the following properties: (i) continuity, (ii) increasingness; i.e., if $x'' >> x' \geq 0_N$, then $F(x'') > F(x')$, (iii) quasiconcavity; i.e., for each utility level u, the upper level set $L(u) \equiv \{x: F(x) \geq u\}$ is convex, (iv) $F(0_N) = 0$ and (v) F(x) tends to $+\infty$ as the components of x all tend to $+\infty$.

[5] C(u,p) is defined for $u \geq 0$, $p >> 0_N$ and has the following properties: (i) it is continuous, (ii) $C(0,p) = 0$ for every $p >> 0_N$, (iii) for every $p >> 0_N$, C(u,p) is increasing in u and C(u,p) tends to $+\infty$ as u tends to $+\infty$, (iv) C(u,p) is positively linearly homogeneous in p for fixed u, i.e., for $u \geq 0$, $p >> 0_N$, $\lambda > 0$, $C(u,\lambda p) = \lambda C(u,p)$, (v) C(u,p) is concave in p for fixed u, (vi) C(u,p) is increasing in p for fixed $u > 0$, i.e., if $p'' >> p' >> 0_N$, $u > 0$, then $C(u,p'') > C(u,p')$ and (vii) C is such that the function $F^*(x) \equiv \max_u \{u: p \cdot x \geq C(u,p) \text{ for every } p >> 0_N\ u \geq 0\}$ is continuous for $x \geq 0_N$.

[6] This is a version of the Shephard [1953] Duality Theorem; see Diewert, [1982].

[7] The term is due to Samuelson [1974].

[8] Throughout this section, we assume that F satisfies Conditions I. Many of the theorems in this section can be proven under much weaker regularity conditions; e.g., see Diewert [1981].

[9] See also Diewert [1973] and Varian [1982].

[10] If $u^0 = u^1$, then $u^* = u^0 = u^1$.

[11] Thus c^r is a flexible functional form to use Diewert's [1974] terminology.

[12] This corresponds to the terminology used in Christensen, Cummings and Jorgenson [1980]. Diewert [1981; p.187] called P_0 the Törnqvist price index, but the term translog price index seems to be more descriptive.

[13] The Allen quantity index is closely related to: (i) Samuelson's [1974] money metric scaling for a consumer's utility function, and (ii) Hicks' [1941-42] **consumer surplus** measures, which are defined in terms of differences of cost functions rather than ratios of cost functions.

[14] The appropriate regularity conditions are listed in Diewert [1982; p.560], and references to the literature on duality theorems between F and D may be found there also. Essentially, D(u,x) has the same regularity properties as the cost function C(u,p) where x replaces p, except that D(u,x) decreases in u while C(u,p) increases in u.

[15] Theorem 7 is not necessarily true if $u > \max\left\{u^0, u^1\right\}$ or if $u < \min\left\{u^0, u^1\right\}$.

[16] Related approximation theorems have been obtained by Samuelson and Swamy [1974] and Vartia [1978]. It should be noted that Diewert's [1978] results were derived using some results due to Vartia [1976].

[17] For example, see Diewert [1978; p.894], Généreux [1983] and Szulc [1983].

[18] In fact Allen and Diewert [1981; p.435] provide an even stronger case for the use of the Fisher ideal formula, since it is the only superlative index number formula that is consistent with both the Hicks [1946; pp. 312-313] and Leontief [1936; pp. 54-57] composite commodity theorems.

[19] We cannot even evaluate how good or bad the approximation is unless we are also given current period quantity information x^t.

[20] Pollak [1975a] demonstrates that it is difficult if not impossible to combine the subindexes into the true cost-of-living index under general conditions on the underlying preferences. This is not exactly the relevant issue, since we cannot calculate the true cost-of-living index anyway in general. However, if we can use the subindexes to form a close approximation to the overall true cost-of-living index, then this is all that we require. Of course, under restrictive assumptions on preferences, subindexes can be combined to give precisely the correct overall cost-of-living index. The first result of this type was obtained by Shephard [1953], who assumed a special structure of preferences that is now called homothetic separability. For generalizations of the Shephard result and reference to the literature, see Blackorby, Primont and Russell [1978; ch. 9].

[21] In Pollak [1975a; p.145], C^m is called a generalized conditional expenditure function for category m.

[22] $C^m(u,p_m,x)$ is non-decreasing in u and non-increasing in the components of x. See McFadden [1978] for a detailed analysis of the properties of C^m.

[23] Remember that $C^m(u,p_m,x)$ does not actually depend on the mth subvector in x, x_m, and hence $P^m(p_m^0,p_m^1,u^*,x^*)$ does not actually depend on x_m^*, the mth subvector in $x^* \equiv (x_1^*,x_2^*,\ldots,x_M^*)$.

[24] Vartia [1976] defines an index number formula to be **consistent in aggregation** if the value of the index calculated in two stages coincides with the value of the index calculated in a single stage. A careful analysis of this concept may be found in Blackorby and Primont [1980]. Empirical evidence on the closeness of two-stage aggregates with the corresponding single-stage aggregates may be found in Diewert [1978, 1983b].

[25] If we are willing to use econometric techniques, then it is not necessary to assume intertemporal additivity in order to estimate the consumer's intertemporal preference function; i.e., see Diewert [1974] and Darrough [1977]. This suggests that it may be possible to adapt the usual index number techniques to the intertemporal context as well.

[26] In the consumer context, see Ruggles [1967], and in the producer context, see Denny and Fuss [1981] and Caves, Christensen and Diewert [1982a].

[27] We are assuming that the household derives disutility from supplying additional hours of work. This may not be true for (x,y) vectors where y is close to 0_M. Formally, we assume that F satisfies Conditions I except that now the domain of definition of F is (x,y): $x \geq 0_N$, $\bar{y} \leq y \leq 0_M$ where $\bar{y} \leq 0_M$.

[28] However, it could still be used as a deflator for the household's nominal "full" income ratio, $(p^1 \cdot x^1 + w^1 \cdot b)/(p^0 \cdot x^0 + w^0 \cdot b)$, to form Pollak implicit quantity indexes as was done in the beginning of Section 3 above. On the concept of "full" income, see Becker [1965].

[29] Recent papers on this issue include McFadyen and Hobart [1978], Rymes [1979], Blinder [1980], Hendershott [1980], Hughes [1980], Gordon [1981], Dougherty and Van Order [1982] and Gillingham [1982].

[30] In fact when one works with user cost formulae of the type defined by (93) and evaluates the expected prices by using *ex post* market prices, one will often find negative user costs for housing.

[31] In technical terms, we replace the original preferences by the convex free disposal hull of the original preferences.

[32] Other types of transactions costs will also have this effect.

[33] Comprehensive accounts of the quality change problem may be found in Triplett [1982] and Hodgins [1982].

* This research was supported by Statistics Canada and the SSHRC of Canada. Neither institution is responsible for the views expressed here. The author is indebted to B. Balk, M. Darrough, David Donaldson, W. Eichhorn, R. Gillingham, B.J. Szulc and J. Weymark for helpful comments.

References

Afriat, S.N. [1967], "The Construction of Utility Functions from Expenditure Data", *International Economic Review* 8, pp.67-77.

_________ [1972], "The Theory of International Comparisons of Real Income and Prices", pp.13-69 in D.J. Daly (ed.), *International Comparisons of Prices and Output*, New York: National Bureau of Economic Research.

_________ [1977], *The Price Index*, London: Cambridge University Press.

Allen, R.G.D. [1949], "The Economic Theory of Index Numbers", *Economica* N.S. 16, pp.197-203.

_________ [1975], "Index Numbers in Theory and Practice", London: Macmillan.

Allen, R.C. and W.E. Diewert [1981], "Direct versus Implicit Superlative Index Number Formulae", *The Review of Economics and Statistics* 63, pp.430-435.

Becker, G.S. [1965], "A Theory of the Allocation of Time", *The Economics Journal*, 75, pp.493-517.

Blackorby, C. and D. Donaldson [1983], "Preference Diversity and Aggregate Economic Cost-of-Living Indexes". *Price Level Measurement: Proceedings From a Conference Sponsored by Statistics Canada.*

_________ D. Primont and R.R. Russell [1978], *Duality Separability and Functional Structure: Theory and Economic Applications*, New York: North-Holland.

_________ and D. Primont [1980], "Index Numbers and Consistency in Aggregation", *Journal of Economic Theory*, 22, pp.87-98.

Blinder, A.S. [1980], ''The Consumer Price Index and the Measurement of Recent Inflation, *Brookings Papers on Economic Activity*, pp.539-565.

Caves, D.W., L.R. Christensen and W.E. Diewert [1982a], ''Multilateral Comparisons of Output, Input and Productivity using Superlative Index Numbers'', *The Economic Journal* 92, pp.73-86.

__________ [1982b], ''The Economic Theory of Index Numbers and the Measurement of Input, Output and Productivity'', *Econometrica* 50, pp.1393-1414.

Christensen, L.R., D. Cummings and D.W. Jorgenson [1980], ''Economic Growth, 1947-73: An International Comparison'', pp.595-691 in *New Developments in Productivity Measurement and Analysism* J.W. Kendrick and B.N. Vaccara (eds.), National Bureau of Economic Research, Studies in Income and Wealth, Vol. 44, Chicago: University of Chicago Press.

Darrough, N.M. [1977],''A Model of Consumption and Leisure in an Intertemporal Framework: A Systematic Treatment using Japanese Data'', *International Economic Review*, 18, 677-696.

Darrough, M. [1983], ''The Treatment of Housing in a Cost of Living Index: Rental Equivalence and User Cost'', *Price Level Measurement: Proceedings From a Conference Sponsored by Statistics Canada.*

Deaton, Angus and J. Muellbauer [1980], *Economics and Consumer Behavior*, London: Cambridge University Press.

Denny, M. [1974], ''The Relationship between Functional Forms for the Production System'', *Canadian Journal of Economics* 7, pp.21-31.

__________ and M. Fuss [1981], ''Intertemporal Changes in Regional Productivity in Canadian Manufacturing,'' *The Canadian Journal of Economics* 14, pp.390-408.

_________ [1983], ''Regional Price Indexes: The Canadian Practice and Some Potential Extensions,'' *Price Level Measurement: Proceedings From a Conference Sponsored by Statistics Canada.*

Diewert, W.E. [1973], ''Afriat and Revealed Preference Theory'', *Review of Economic Studies* 40, pp.419-425.

_________ [1974], ''Intertemporal Consumer Theory and the Demand for Durables'', *Econometrica* 42, pp.497-516.

_________ [1976], ''Exact and Superlative Index Numbers'', *Journal of Econometrics* 4, pp.115-145.

_________ [1978], ''Superlative Index Numbers and Consistency in Aggregation'', *Econometrica* 46, pp.883-900.

_________ [1980], ''Aggregation Problems in the Measurement of Capital'', pp.433-528 in *The Measurement of Capital*, D. Usher (ed.), Chicago: The University of Chicago Press.

_________ [1981], ''The Economic Theory of Index Numbers: A Survey,'' pp.163-208 in *Essays in the Theory and Measurement of Consumer Behavior in Honour of Sir Richard Stone*, Angus Deaton, (ed.), Cambridge, England: Cambridge University Press.

_________ [1982a], ''Duality Approaches to Microeconomic Theory'', pp.535-599 in *Handbook of Mathematical Economics*, Vol. II, K.J. Arrow and M.D. Intriligator, (eds.), Amsterdam: North-Holland.

_________ [1983b], ''The Treatment of Seasonality in a Cost-of-Living Index'', *Price Level Measurement: Proceedings From a Conference Sponsored by Statistics Canada.*

_________ [1983], ''The Theory of the Output Price Index and the Measurement of Real Output Change'', *Price Level Measurement: Proceedings From a Conference Sponsored by Statistics Canada.*

Dougherty, Ann and R. Van Order [1982], ''Inflation, Housing Costs and the Consumer Price Index'', *American Economic Review* 72, pp.154-164.

Eichhorn, W. [1976], ''Fisher's Tests Revisited'', *Econometrica* 44, pp.247-256.

_________ [1978], *Functional Equations in Economics*, Reading, Mass.: Addison-Wesley.

_________ and J. Voeller [1976], *Theory of the Price Index: Fisher's Test Approach and Generalizations*, Lecture Notes in Economics and Mathematical Systems, Berlin: Springer-Verlag.

_________ [1983], ''The Axiomatic Foundations of Price Indexes and Purchasing Power Parities'', *Price Level Measurement: Proceedings From a Conference Sponsored by Statistics Canada.*

Fisher, Irving [1922], *The Making of Index Numbers*, Boston: Houghton-Mifflin.

Généreux, P. [1983], ''Impact of Alternative Formulae on the CPI'', *Price Level Measurement: Proceedings From a Conference Sponsored by Statistics Canada.*

Gillingham, G. [1982], ''Measuring the Cost of Shelter for Homeowners: Theoretical and Empirical Considerations'', Division of Price and Index Number Research, Office of Prices and Living Conditions, BLS, Washington, D.C.

Gordon, R.J. [1981], ''The Consumer Price Index: Measuring Inflation and Causing It'', *The Public Interest* 63, pp.112-134.

Hendershott, P. [1980], ''Real User Costs and the Demand for Single-Family Housing'', *Brookings Papers on Economic Activity*, pp.401-444.

Hicks, J.R. [1940], "The Valuation of the Social Income", *Economica* 7, pp.105-124.

________ [1941-42], "Consumers' Surplus and Index Numbers", *Review of Economic Studies*, 9, pp.126-137.

________ [1946], *Value and Capital*, second edition, Oxford: Clarendon Press.

________ [1958], "The Measurement of Real Income", *Oxford Economic Papers*, 10, pp. 125-162.

________ [1961], "Measurement of Capital in Relation to the Measurement of other Economic Aggregates", in F.A. Lutz and D.C. Hague (eds.), *The Theory of Capital, London*: Macmillan.

________ [1981], *Wealth and Welfare*, Cambridge, Mass.: Harvard University Press.

Hodgins, C.D. [1982], "Inflation Adjustments and Indexation: Is the Consumer Price Index an Appropriate Index?", mimeo, September 30.

Hughes, G.A. [1980], "Housing and the Tax System", pp.67-105, in *Public Policy and the Tax System*, G.A. Hughes, and G.M. Heal (eds.), London: George Allen and Unwin.

Jorgenson, D.W. and D.T. Slesnick [1983], "Individual and Social Cost-of-Living Indexes", *Price Level Measurement: Proceedings From a Conference Sponsored by Statistics Canada.*

Joseph, M.F.W. [1935-36], "Mr. Lerner's Supplementary Limits for Price Index Numbers", *Review of Economic Studies*, 3, pp.155-157.

Konüs, A.A. [1924], "The Problem of the True Index of the Cost of Living," translated in *Econometrica* 7, 1939, pp.10-29.

Leontief, W. [1936], "Composite Commodities and the Problem of Index Numbers", *Econometrica* 4, pp.39-59.

Lerner, A.P. [1935-36], ''A Note on the Theory of Price Index Numbers'', *Review of Economic Studies* 3, pp.50-56.

Loyns, R.M.A. [1972], *An Examination of the Consumer Price Index and Implicit Price Index as Measures of Recent Price Changes in the Canadian Economy*, Ottawa: Information Canada.

Malmquist, S. [1953], ''Index Numbers and Indifference Surface'', *Trabajos de Estadistica* 4, pp.209-242.

McFadden, D. [1978], ''Cost, Revenue and Profit Functions'', pp.3-109 in *Production Economics; A Dual Approach to Theory and Applications*, Vol. 1, M. Fuss and D. McFadden (eds.), Amsterdam: North-Holland.

McFadyen, S. and R. Hobart [1978], ''An Alternative Measurement of Housing Costs and the Consumer Price Index'', *Canadian Journal of Economics* 11, pp.105-112.

Muellbauer, J. [1974], ''The Political Economy of Price Indices'', Birkbeck Discussion Paper 22, March.

Pigou, A.C. [1920], *The Economics of Welfare*, London: Macmillan.

Pollak, R.A. [1969], ''Conditional Demand Functions and Consumption Theory'', *Quarterly Journal of Economics*, 83, pp.60-78.

__________ [1983], ''The Theory of the Cost-of-Living Index,'' *Price Level Measurement: Proceedings From a Conference Sponsored by Statistics Canada.*

__________ [1975a], ''Subindexes of the Cost-of-Living Index'', *International Economic Review* 16, pp.135-150.

_________ [1975b], "The Intertemporal Cost-of-Living Index", *Annals of Economic and Social Measurement* 4, pp.179-196.

_________ [1981], "The Social Cost-of-Living Index", *Journal of Public Economics* 15, pp.311-336.

Prais, S. [1959], "Whose Cost-of-Living?", *Review of Economic Studies* 26, pp.126-134.

Riddell, W.C. [1983], "The Consumer Price Index and the Value of Leisure Time' *Price Level Measurement: Proceedings From a Conference Sponsored by Statistics Canada.*

Ruggles, R. [1967], "Price Indexes and International Price Comparisons", pp.171-205 in *Ten Economic Studies in The Tradition of Irving Fisher*, W. Fellner et al (eds.), New York: John Wiley.

Rymes, T.K. [1979], "The Treatment of Home Ownership in the CPI", *Review of Income and Wealth* 25, pp.393-412.

Samuelson, P.A. [1947], *Foundations of Economic Analysis*, Cambridge, Mass.: Harvard University Press.

_________ [1950], "Evaluation of Real National Income", *Oxford Economic Papers* 2, pp.1-29.

_________ [1974], "Complementarity - An Essay on the 40th Anniversary of the Hicks-Allen Revolution in Demand Theory", *The Journal of Economic Literature* 12, pp.1255-1289.

_________ and S. Swamy [1974], "Invariant Economic Index Numbers and Canonical Duality: Survey and Synthesis", *American Economic Review*, 64, pp.566-593.

Szulc, B.J. [1982], "Linking Index Number Series", *Price Level Measurement: Proceedings From a Conference Sponsored by Statistics Canada.*

Shephard, R.W. [1953], *Cost and Production Functions*, Princeton: Princeton University Press.

Statistics Canada [1982], ''The Consumer Price Index Reference Paper: Concepts and Procedures'', Catalogue No. 62-553, Occasional, Ottawa, Canada.

Triplett, J.E. [1982], ''Concepts of Quality in Input and Output Price Measures: A Resolution of the User Value-Resource Cost Debate'', in *The U.S. National Income and Product Accounts: Selected Topics*, M. Fuss (ed.), Chicago, The University of Chicago Press.

Usher, D. [1980], *The Measurement of Economic Growth*, New York; Columbia University Press.

Varian, H. [1982], ''The Nonparametric Approach to Demand Analysis'', *Econometrica*, 50, pp.945-973.

Vartia, Y.O. [1976], ''Ideal Log-Change Index Numbers'', *Scandinavian Journal of Statistics* 3, pp.121-126.

__________ [1978], ''Fisher's Five Tined Fork and other Quantum Theories of Index Numbers'', pp.271-295 in *Theory and Economic Applications of Economic Indices*, W. Eichhorn, R. Henn, O. Opitz and R.W. Shephard, (eds.), Wurzburg: Physica-Verlag.

COMMENTS

R. Robert Russell
Department of Economics
New York University

Selling an economics paper is much like selling a consumer product; its sales prospects depend on packaging and promotion as well as content. My own casual empirical investigation suggests that there might even be an inverse correlation between content and promotion, much as inferior chiantis come in fancy straw baskets while the finer chiantis come in plain Bordeaux-type bottles.

When it comes to the ratio of packaging to content, Erwin Diewert's papers are at the extreme left-hand side of the spectrum – not only because the content is always so substantial, but also because he so eschews promotion.

Take, for example, Diewert's Theorem 1. To characterize this theorem, I draw your attention to the rather bifurcated research on index number theory in recent years. In the so-called mechanistic approach, represented at this conference in the paper by Wolfgang Eichhorn [1983], index numbers are not based on utility maximization; indeed, no assumptions are made about relationships between prices and quantities (i.e., about demand curves). The alternative approach, based on the theory of utility maximization, has a long history, outlined in many of the papers presented at this conference (especially Pollak [1983]). By providing necessary and sufficient conditions for a mechanistic price index to be rationalized by the theory of constrained cost minimization, Diewert's theorem synthesizes the two seemingly unrelated approaches to cost-of-living indexes. So how does Diewert characterize his results? As follows: "The above theorem is very closely related to some results of Pollak [1983] ...". Diewert's result is closely related to that of Pollak, but it is also much more.

All of Diewert's papers are densely packed with valuable results, and this one is no exception. This paper also evinces another important aspect of his papers: their high scholarly quality – not just, or even primarily, mathematical rigor, but also the careful attention that is paid to the history of economic thought. To a certain extent, this paper is a rigorous exercise in the history of economic thought, since many of the results are traced to the

early work of, among others, Konüs and Malmquist. Especially important is the Konüs result on the existence of a reference utility level that bounds the true cost-of-living index between the Paasche and Laspeyres indexes. This result, which is used extensively by Diewert throughout this paper, is especially important when the Paasche and Laspeyres indexes are "close" to one another, in which case the true cost-of-living index for the intermediate reference utility level is fairly narrowly delineated. As Diewert puts it, "We will have the consumer's true cost-of-living between periods 0 and 1 for all practical purposes".

But enough accolades. In the remainder of this comment, I discuss one aspect of the paper that I believe is expositionally deficient. In particular, I attempt to shed some light on the relationship between two classes of indexes examined in Diewert's paper: The Konüs price index (the usual utility-based formula) and the Malmquist quantity index (which Diewert, with characteristically good taste, seems to favour among the several quantity indexes that he discusses).

In Diewert's presentation, these two indexes seem to be quite different types of animals – one based on constrained cost minimization and the other on distance functions. In fact, because of some important duality results that can be found in Shephard [1970] and Blackorby, Primont, and Russell [1978, Ch. 2], the Konüs and Malmquist indexes are completely symmetric and constitute a natural pair of quantity and price indexes. (This symmetry was first pointed out in Blackorby and Russell [1977].)

Diewert's (standard) derivation of the Konüs cost-of-living index is summarized in panel A of Table 1. The cost function, C, is derived by choosing a consumption vector, x, to minimize total expenditure subject to a utility constraint. The value, C(u,p), can be written as the inner product of the price vector, p, and the vector-valued, constant-utility (Hicksian) demand function, $\zeta(u,p)$. This minimization problem is depicted in the diagram by the tangency of the equal-cost line (with normal $|\bar{p}|$) and the u indifference curve. The value of the cost function at $(\bar{u},\bar{p})$ is given by the horizontal intercept of the equal-cost line multiplied by $\bar{p}_1$. The Konüs cost-of-living index is then defined as the ratio of the values of the cost functions for the two different price vectors, p^0 and p^1.

TABLE 1

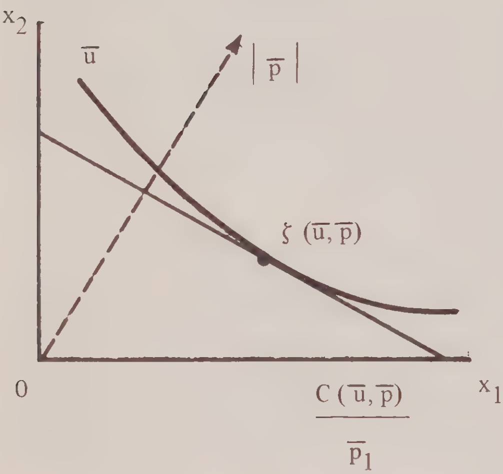
THE KONUS AND MALMQUIST INDICES CONSTRUCTED WITH COST FUNCTIONS

A

Derivation of Konus
Cost-of-Living
(Price) Index

$$C(u,p) = \min_{x} \{p\cdot x \mid U(x) \geq u\}$$

$$= p\cdot\zeta(u,p)$$



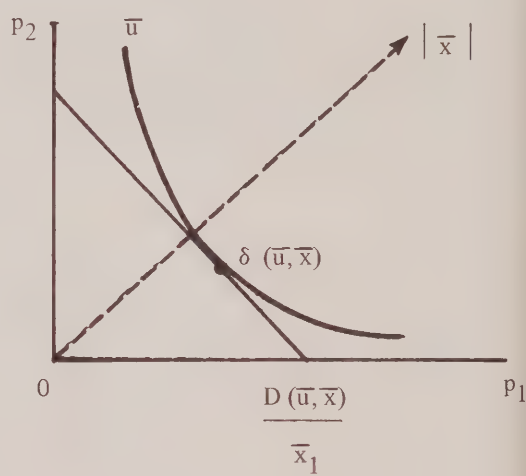
$$P_K(p^0,p^1,u) = \frac{C(u,p^1)}{C(u,p^0)}$$

B

Derivation of Malmquist
Standard-of-Living
(Quantity) Index

$$D(u,x) = \min_{p} \{p\cdot x \mid V(p) \leq u\}$$

$$= \delta(u,x)\cdot x$$



$$Q_M(x^0,x^1,u) = \frac{D(u,x^1)}{D(u,x^0)}$$

Diewert's derivation of the Malmquist quantity index (called the standard-of-living index by Blackorby and Russell [1977]), is shown in panel B of Table 2. The distance function, D, with image D(u,x), is defined as the maximal (proportional) amount by which x can be reduced and still be contained in the u upper level set. Thus, in the diagram, the distance function for utility level $\bar{u}$ and consumption vector $\bar{x}$ is given by the ratio α/β. The Malmquist quantity index is defined as the ratio of these distance functions for two different quantity vectors, x^0 and x^1.

These two concepts and their derivations seem quite disparate. Because of the duality

between distance functions and cost functions, however, they are perfectly symmetric. The Malmquist quantity index can alternatively be derived by the cost-minimization procedure shown in panel B of Table 1. Here, the function D is defined as the minimum imputed value of a given quantity vector, subject to the constraint that the indirect utility function evaluated at the chosen (shadow) price vector be no greater than u. This (imputed) cost function is identical to the distance function in panel B of Table 2. It can be written as the inner product of the vector-valued, constant-utility imputed-price function, $\delta(u,x)$, and the quantity vector, x.

The cost-minimization problem is illustrated in panel B by the tangency between the plane with normal $|\bar{x}|$ and the indirect indifference curve labelled $\bar{u}$ at the price vector $\delta(\bar{u},\bar{x})$. The function value, $D(\bar{u},\bar{x})$, can then be obtained by multiplying the horizontal intercept of this plane by $\bar{x}_1$. The Malmquist quantity index is then defined as the ratio of values of the cost-imputation function at two different quantity vectors, x^0 and x^1. The juxtaposition of panels A and B of Table 1 make the complete symmetry between this derivation of the Malmquist quantity index and the standard derivation of the Konüs cost-of-living index apparent. The interchange of $(x,U,C,\geq)$ and $(p,V,D,\leq)$ yields equivalent derivations.

Just as the Malmquist index, defined as the ratio of distance functions by Diewert, can also be written as a ratio of cost-imputation functions, so can the Konüs cost-of-living index, traditionally defined as the ratio of cost functions, be defined as a ratio of distance functions. This is illustrated in panel A of Table 2. $C(u,p)$ is the maximum proportional amount by which the price vector, p, can be reduced while placing it in the u lower-level set in price space (recall that V is non-increasing in p). In the diagram, $C(\bar{u},\bar{p})$ is given by the ratio of the two line segments in the diagram, α/β. The Konüs cost-of-living index can be written as the ratio of values of the distance functions for two different price vectors, p^0 and p^1. The complete symmetry between these two indexes is revealed by comparing panels A and B of Table 2, where each is depicted as a ratio of values of a distance function. The interchange of $(x,U,D,\geq)$ and $(p,V,C,\leq)$ yields equivalent derivations.

To conclude, the Malmquist quantity index, which Diewert favours because of its intrinsic properties, is in fact the *only* natural counterpart to the widely accepted Konüs cost-of-living index.

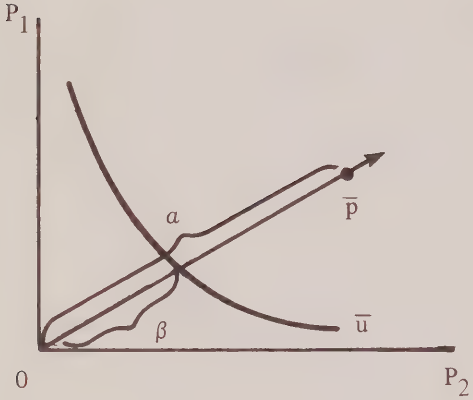
TABLE 2

THE KONUS AND MALMQUIST INDICES CONSTRUCTED WITH DISTANCE FUNCTIONS

A

Derivation of Konus
Cost-of-Living
(Price) Index

$$C(u,p) = \max_{\lambda} \{\lambda \mid V(p/\lambda) \leq u\}$$



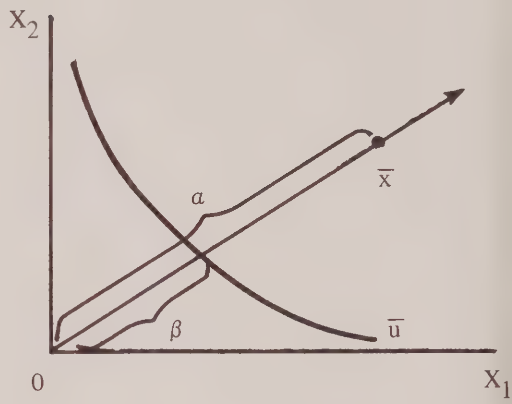
$$C(\bar{u}, \bar{p}) = a/\beta$$

$$P_K(p^0, p^1, u) = \frac{C(u, p^1)}{C(u, p^0)}$$

B

Derivation of Malmquist
Standard-of-Living
(Quantity) Index

$$D(u,x) = \max_{\lambda} \{\lambda \mid U(x/\lambda) \geq u\}$$



$$D(\bar{u}, \bar{x}) = a/\beta$$

$$Q_M(x^0, x^1, u) = \frac{D(u, x^1)}{D(u, x^0)}$$

References

Blackorby, C., D. Primont and R.R. Russell [1978], *Duality Separability, and Functional Structure; Theory and Economic Applications*, New York: North-Holland.

_________ and R.R. Russell [1977], "Indices and Subindices of the Cost of Living and the Standard of Living", *International Economic Review* 18, pp.229-240.

Eichhorn, W. [1983], "The Axiomatic Foundations of Price Indices and Purchasing Power Parities", *Price Level Measurement: Proceedings From a Conference Sponsored by Statistics Canada.*

Pollak, R.A. [1983], "The Theory of the Cost-of-Living Index", *Price Level Measurement: Proceedings From a Conference Sponsored by Statistics Canada.*

Shephard, R. [1970], *The Theory of Cost and Production Functions*, Princeton: Princeton University Press.

INDIVIDUAL AND SOCIAL COST-OF-LIVING INDEXES

Dale W. Jorgenson
Department of Economics
Harvard University
Daniel T. Slesnick
Department of Economics
University of Texas

SUMMARY

The purpose of this paper is to present an econometric approach to cost-of-living measurement. This approach implements the economic theory of individual cost-of-living measurement pioneered by Konüs [1939] almost six decades ago. In this paper we develop and implement a completely parallel theory of social cost-of-living measurement.

Our approach to cost-of-living measurement is based on an econometric model of aggregate consumer behavior. The novel feature of this model is that systems of individual demand functions can be recovered uniquely from the system of aggregate demand functions. We derive cost-of-living indexes for individual households from systems of individual demand functions.

Our key innovation in the economic theory of cost-of-living measurement is the introduction of an explicit social welfare function. Our social welfare function incorporates measures of individual welfare from our econometric model. In addition, this social welfare function employs normative criteria for evaluating transfers among individuals.

Given measures of individual welfare from our econometric model, we can express the level of social welfare as a function of prices and of total expenditures of all consuming units. We present methods for translating changes in prices into measures of change in the social cost-of-living. Our definitions of social and individual cost-of-living indexes are perfectly analogous.

Finally, we extend the concept of a social cost-of-living index to groups of consuming units with common demographic characteristics. Our definition of the group cost-of-living index is analogous to the definition of a social cost-of-living index. To implement a group cost-of-living index we require a group welfare function that is analogous to a social welfare function.

RÉSUMÉ

Le présent document vise à exposer un modèle d'approche économétrique du calcul du coût de la vie. Cette approche fait appel à la théorie économique du calcul individuel du coût de la vie dont Konüs [1939] a été le pionnier il y a près de 60 ans.[1] *Le présent document élabore et met en oeuvre une théorie tout à fait parallèle du calcul du coût social de la vie.*

Notre approche du calcul du coût de la vie repose sur un modèle économétrique du comportement global des consommateurs. La nouveauté de ce système réside dans le fait que les systèmes de fonctions de demande individuelle ne peuvent être récupérés qu'à partir du système de fonctions de demande globale. Nous établissons les indices du coût de la vie pour des ménages à partir de systèmes de fonctions individuelles de demande.

L'introduction d'une fonction explicite de bien-être social constitue l'innovation clé au niveau de la théorie économique du calcul du coût de la vie. Notre fonction de bien-être social intègre des calculs de bien-être individuel a partir de notre modèle économétrique. Cette fonction de bien-être social fait également appel à des critères normatifs d'évaluation des transferts entre individus.

Dans la partie 2, nous exposons la méthodologie économétrique servant à mettre au point un modèle de comportement global des consommateurs. Dans ce modèle, le système de fonctions de demande globale repose sur des statistiques sommaires de la distribution conjointe des attributs et des dépenses totales des ménages. Les attributs des ménages tels que les caractéristiques démographiques nous permettent de rendre compte des écarts au niveau des préférences.

Dans la partie 3, nous appliquons notre modèle écnométrique de comportement global des consommateurs aux États-Unis. À cette fin, nous utilisons des données transversales sur les régimes individuels de dépenses. Nous combinons ces données aux séries chronologiques sur les régimes de dépenses globales. Nous employons également des données des séries chronologiques sur la répartition des dépenses totales entre unités de consommation.

Dans la partie 4, nous exposons les méthodes servant à convertir les variations de prix en mesures de variation du coût individuel de la vie. À cette fin, nous employons la fonction individuelle de dépense. La fonction de dépense permet d'établir les dépenses minimales totales nécessaires pour atteindre un niveau de base de bien-être individuel. Ces dépenses minimales sont fonction des prix et des attributs de l'unité de consommation.

Suivant Konüs [1939], nous définissons l'indice du coût individuel de la vie comme étant le rapport entre les dépenses totales nécessaires pour atteindre un niveau de base de bien-être individuel aux prix courants et le niveau de base correspondant des dépenses. Lorsque l'indice du coût individuel de la vie excède l'unité et que les dépenses totales sont constantes, le bien-être de l'unité de consommation a diminué par rapport au niveau de base.

Le problème qui reste, lorsque l'on veut établir l'indice du coût individuel de la vie, c'est de déterminer le niveau de base de bien-être individuel. À cette fin, nous définissons le bien-être individuel en termes de fonctions d'utilité indirectes. Cette fonction exprime l'utilité comme une fonction des prix de même que des dépenses totales et des attributs de l'unité de consommation.

Nous calculons les indices du coût individuel de la vie et les taux d'inflation en fonction de ménages avec différents niveaux de dépenses totales et diverses caractéristiques démographiques pour la période comprise entre 1958 et 1978. À cette fin, nous établissons des fonctions d'utilité indirecte et des fonctions de dépenses pour toutes les unités de consommation à partir de notre modèle économétrique de comportement global des consommateurs.

Dans la partie 4, nous énumérons les méthodes servant à évaluer le niveau de bien-être social. À cette fin, nous établissons une fonction explicite de bien-être social. Notre fonction de bien-être social intègre les calculs du bien-être individuel fondés sur des fonctions

d'utilité indirecte de toutes les unités de consommation. Cette fonction de bien-être social fait également appel à des critères normatifs qui reposent sur les valeurs horizontales et verticales servant à évaluer les transferts entre unités.

Une fois établies les fonctions d'utilité indirecte à partir de notre modèle économétrique, nous pouvons exprimer le niveau de bien-être social en tant que fonction des prix de même que des dépenses totales et des attributs de toutes les unités de consommation. Nous définissons la fonction dépenses sociales comme étant les dépenses minimales globales nécessaires pour atteindre un niveau de base de bien-être social. Ce niveau minimal de dépenses est fonction des prix et des attributs de chacune des unités de consommation.

Dans la partie 6, nous exposons les méthodes servant à convertir les modifications au niveau des prix en mesures de changement du coût social de la vie. Suivant Pollak [1981], nous définissons l'indice du coût social de la vie comme étant le rapport entre les dépenses globales nécessaires pour atteindre un niveau de base de bien-être social aux prix courants et le niveau de base correspondant de dépenses. Lorsque l'indice du coût social de la vie excède l'unité et que les dépenses globales sont constantes, le bien-être social a diminué relativement au niveau de base.

Nos définitions d'indices de coûts social et individuel de la vie sont parfaitement analogues. Dans ces définitions, les rôles des fonction de dépenses sociales et individuelles et ceux des fonctions de bien-être social et individuel sont parfaitement parallèles. Nous calculons l'indice du coût social de la vie et les taux d'inflation pour la période comprise entre 1958 et 1978.

Dans la partie 7, nous appliquons le concept d'indice du coût social de la vie à des groupes d'unités de consommation possédant des caractéristiques démographiques communes. Notre définition de l'indice du coût de groupe de la vie est analogue à celle de l'indice du coût social de la vie. Pour établir un indice du coût de groupe de la vie, nous avons besoin d'une fonction de bien-être de groupe analogue à une fonction de bien-être social. Nous avons également besoin d'une fonction de dépenses de groupe.

Une fois les fonctions de bien-être et de dépenses de groupe établies, nous calculons les

indices de coût de la vie et les taux d'inflation par rapport à des groupes de ménages possédant des caractéristiques démographiques communes pour la période s'échelonnant entre 1958 et 1978. Lorsque l'indice du coût de groupe de la vie dépasse l'unité et que les dépenses de groupe restent constantes, le bien-être du groupe a diminué relativement au niveau de base.

À la partie 8, nous comparons les approches économétrique et indicielle au calcul du coût de la vie. L'approche économétrique intègre toutes les données servant à établir les indices du coût de la vie. L'approche économétrique, toujours, permettrait, entre autres avantages majeurs, de résumer les données disponibles de façon concise et facilement intelligible.

En conclusion, nous soulignons que les approches économétrique et indicielle ont toutes les deux d'importantes limites, puisqu'elles posent certains problèmes pratiques lorsque l'on veut obtenir des données appropriées sur les prix et les dépenses. L'approche économétrique est cependant plus souple et plus facile à appliquer. L'établissement des indices du coût individuel, social et de groupe de la vie pour les États-Unis et pour la période comprise entre 1958 et 1978 illustre ces avantages.

1. Introduction

The purpose of this paper is to present an econometric approach to cost-of-living measurement. This approach implements the economic theory of individual cost-of-living measurement pioneered by Konüs [1939] almost six decades ago.[1] In this paper we develop and implement a completely parallel theory of social cost-of-living measurement.

Our approach to cost-of-living measurement is based on an econometric model of aggregate consumer behavior. The novel feature of this model is that systems of individual demand functions can be recovered uniquely from the system of aggregate demand functions. We derive cost-of-living indexes for individual households from systems of individual demand functions.

Our key innovation in the economic theory of cost-of-living measurement is the introduction of an explicit social welfare function. Our social welfare function incorporates measures of individual welfare from our econometric model. In addition, this social welfare function employs normative criteria for evaluating transfers among individuals.

In Section 2 we outline econometric methodology for developing a model of aggregate consumer behavior. In this model the system of aggregate demand functions depends on summary statistics of the joint distribution of attributes and total expenditures of individual households. Attributes of households such as demographic characteristics enable us to account for differences in preferences.

In Section 3 we implement our econometric model of aggregate consumer behavior for the United States. For this purpose we employ cross-section data on individual expenditure patterns. We combine these data with time series information on aggregate expenditure patterns. We also employ time series data on the distribution of total expenditures among consuming units.

In Section 4 we present methods for translating changes in prices into measures of change in the individual cost of living. For this purpose we employ the individual expenditure function. The expenditure function gives the minimum total expenditure required to attain a base level of individual welfare. This minimum expenditure depends on prices and on the attributes of the consuming unit.

Following Konüs [1939], we define the individual cost-of-living index as the ratio between the total expenditure required to attain a base level of individual welfare at current prices and the corresponding base level of expenditure. If the individual cost-of-living index exceeds unity and total expenditure is constant, then the welfare of the consuming unit has decreased relative to the base level.

To implement the individual cost-of-living index the remaining problem is to determine the base level of individual welfare. For this purpose we define individual welfare in terms of the indirect utility function. This function gives utility as a function of prices and of the total expenditure and attributes of the consuming unit.

We calculate individual cost-of-living indexes and rates of inflation for households with different levels of total expenditure and different demographic characteristics for the period 1958-1978. For this purpose we derive indirect utility functions and expenditure functions for all consuming units from our econometric model of aggregate consumer behavior.

In Section 5 we present methods for evaluating the level of social welfare. For this purpose we construct an explicit social welfare function. Our social welfare function incorporates measures of individual welfare based on indirect utility functions for all consuming units. In addition, this social welfare function employs normative criteria based on horizontal and vertical equity for evaluating transfers among units.

Given indirect utility functions from our econometric model, we can express the level of social welfare as a function of prices and of total expenditures and attributes for all consuming units. We define the social expenditure function as the minimum aggregate expenditure required to attain a base level of social welfare. This minimum level of expenditure depends on prices and on the attributes of all consuming units.

In Section 6 we present methods for translating changes in prices into measures of change in the social cost of living. Following Pollak [1981], we define the social cost-of-living index as the ratio between the aggregate expenditure required to attain a base level of social welfare at current prices and the corresponding base level of expenditure. If the social cost-of-living index exceeds unity and aggregate expenditure is constant, then social welfare has decreased relative to the base level.

Our definitions of social and individual cost-of-living indexes are perfectly analogous. In these definitions the roles of the social and individual expenditure functions and the roles of the social and individual welfare functions are completely parallel. We calculate the social cost-of-living index and rates of inflation for the period 1958-1978.

In Section 7 we extend the concept of a social cost-of-living index to groups of consuming units with common demographic characteristics. Our definition of the group cost-of-living index is analogous to the definition of a social cost-of-living index. To implement a group cost-of-living index we require a group welfare function that is analogous to a social welfare function. We also require a group expenditure function.

Given group welfare and expenditure functions, we calculate cost-of- living indexes and rates of inflation for groups of households with common demographic characteristics for the period 1958-1978. If a group cost-of- living index exceeds unity and group expenditure is constant, then the welfare of the group has decreased relative to the base level.

In Section 8 we compare the econometric and index number approaches to cost-of-living measurement. The econometric approach incorporates all the information employed in cost-of-living index numbers. An important advantage of the econometric approach is that it summarizes the available information in a concise and readily intelligible way.

In concluding we emphasize that the econometric and index number approaches share a number of significant limitations. These limitations arise from the practical problems of obtaining appropriate data on prices and expenditures. However, the econometric approach has greater flexibility and is easier to apply. These advantages are illustrated by our implementation of individual, social and group cost-of-living indexes for the United States for the period 1958-1978.

2. Aggregate Consumer Behavior

In this section we develop an econometric model of aggregate consumer behavior based on the theory of exact aggregation, following Jorgenson, Lau and Stoker [1980, 1981, 1982]. Our model incorporates time series data on prices and aggregate quantities consumed. We also include cross-section data on individual quantities consumed, individual total expenditure, and attributes of individual households such as demographic characteristics.

To construct an econometric model based on exact aggregation we first represent individual preferences by means of an indirect utility function for each consuming unit, using the following notation:

p_n – price of the nth commodity, assumed to be the same for all consuming units.

$p = (p_1, p_2 \dots p_N)$ – the vector of prices of all commodities.

x_{nk} – the quantity of the nth commodity group consumed by the kth consuming unit ($n = 1, 2 \ldots N$; $k = 1, 2 \ldots K$).

$M_k = \sum_{n=1}^{N} p_n x_{nk}$ – total expenditure of the kth consuming unit ($k = 1, 2 \ldots K$).

$w_k = p_n x_{nk} / M_k$ – expenditure share of the nth commodity group in the budget of the kth consuming unit ($n = 1, 2 \ldots N$; $k = 1, 2 \ldots K$).

$w_k = (w_{1k}, w_{2k} \cdots w_{Nk})$ – vector of expenditure shares for the kth consuming unit ($k = 1, 2 \ldots K$).

$\ln \frac{p}{M_k} = (\ln \frac{p_1}{M_k}, \ln \frac{p_2}{M_k} \ldots \ln \frac{p_N}{M_k})$ – vector of logarithms of ratios of prices to expenditure by the kth consuming unit ($k = 1, 2 \ldots K$).

$\ln p = (\ln p_1, \ln p_2 \ldots \ln p_N)$ – vector of logarithms of prices.

A_k – vector of attributes of the kth consuming unit ($k = 1, 2 \ldots K$).

We assume that the kth consuming unit allocates expenditures in accord with the transcendental logarithmic or translog indirect utility function,[2] say V_k, where:

$$\ln V_k = G(\ln \frac{p}{M_k}{}'\alpha_p + \frac{1}{2} \ln \frac{p}{M_k}{}' B_{pp} \ln \frac{p}{M_k} + \ln \frac{p}{M_k}{}' B_{pA} A_k, A_k), \quad (k = 1, 2 \ldots K). \tag{2.1}$$

In this representation the function G is a monotone increasing function of the variable $\ln \frac{p}{M_k}{}'\alpha_p + \frac{1}{2} \ln \frac{p}{M_k}{}' B_{pp} \ln \frac{p}{M_k} + \ln \frac{p}{M_k}{}' B_{pA} A_k$. In addition, the function G depends directly on the attribute vector A_k.[3] The vector α_p and the matrices B_{pp} and B_{pA} are constant parameters that are the same for all consuming units.

The expenditure shares of the kth consuming unit can be derived by the logarithmic form

of Roy's [1943] Identity:[4]

$$w_{nk} = \frac{\partial \ln V_k}{\partial \ln (p_n/M_k)} \Big/ \sum_{n=1}^{N} \frac{\partial \ln V_k}{\partial \ln (p_n/M_k)}, \quad (n = 1, 2 \ldots N;\ k = 1, 2 \ldots K). \tag{2.2}$$

Applying this Identity to the translog indirect utility function (2.1), we obtain the system of individual expenditure shares:

$$w_k = \frac{1}{D_k(p)} (\alpha_p + B_{pp} \ln \frac{p}{M_k} + B_{pA} A_k), \quad (k = 1, 2 \ldots K), \tag{2.3}$$

where the denominators $\{D_k\}$ take the form:

$$D_k = i'\alpha_p + i'B_{pp} \ln \frac{p}{M_k} + i'B_{pA} A_k, \quad (k = 1, 2 \ldots K). \tag{2.4}$$

The individual expenditure shares are homogeneous of degree zero in the unknown parameters -- α_p, B_{pp}, B_{pA}. By multiplying a given set of these parameters by a constant we obtain another set of parameters that generates the same system of individual budget shares. Accordingly, we can choose a normalization for the parameters without affecting observed patterns of individual expenditure allocation. We find it convenient to employ the normalization:

$$i'\alpha_p = -1 .$$

Under this restriction any change in the set of unknown parameters will be reflected in changes in individual expenditure patterns.

The conditions for exact aggregation are that the individual expenditure shares are linear in functions of the attributes $\{A_k\}$ and total expenditures $\{M_k\}$ for all consuming units[5]. These conditions will be satisfied if and only if the terms involving the attributes

and expenditures do not appear in the denominators of the expressions given above for the individual expenditure shares, so that:

$$i'B_{pp}i = 0 ,$$

$$i'B_{pA} = 0 .$$

The exact aggregation restrictions imply that the denominators $\{D_k\}$ reduce to:

$$D = -1 + i'B_{pp}\ln p,$$

where the subscript k is no longer required, since the denominator is the same for all consuming units. Under these restrictions the individual expenditure shares can be written:

$$w_k = \frac{1}{D(p)}(\alpha_p + B_{pp}\ln p - B_{pp}\, i \cdot \ln M_k + B_{pA}A_k), \quad (k = 1, 2, \ldots K). \tag{2.5}$$

The individual expenditure shares are linear in the logarithms of expenditures $\{\ln M_k\}$ and in the attributes $\{A_k\}$, as required by exact aggregation.

Under exact aggregation the indirect utility function for each consuming unit can be represented in the form:

$$\ln V_k = F(A_k) + \ln p'(\alpha_p + \frac{1}{2} B_{pp}\ln p + B_{pA}A_k) - D(p) \ln M_k , \quad (k = 1, 2, \ldots K). \tag{2.6}$$

In this representation the indirect utility function is linear in the logarithm of total expenditure $\ln M_k$ with a coefficient that depends on the prices p (k = 1, 2 ... K). This property is invariant with respect to positive affine transformations, but is not preserved by arbitrary monotone increasing transformations. We conclude that the indirect utility function (2.6) provides a cardinal measure of utility for each consuming unit.

If a system of individual expenditure shares (2.3) can be generated from an indirect utility

function of the form (2.1) we say that the system is **integrable**. A complete set of conditions for integrability[6] is the following:

1. Homogeneity. The individual expenditure shares are homogeneous of degree zero in prices and total expenditure.

We can also write the individual expenditure shares in the form:

$$\beta_{pM} = B_{pp} i. \tag{2.7}$$

Given the exact aggregation restrictions, there are N-1 restrictions implied by homogeneity.

2. Summability. The sum of the individual expenditure shares over all commodity groups is equal to unity:

$$i' w_k = 1, \qquad (k = 1, 2 \ldots K).$$

We can write the denominator D(p) in (2.4) in the form:

$$D = -1 + \beta_{Mp} \ln p,$$

where the vector of parameters β_{Mp} is constant and the same for all commodity groups and all consuming units. Summability implies that this vector must satisfy all restrictions:

$$\beta_{Mp} = i' B_{pp} . \tag{2.8}$$

Given the exact aggregation restrictions, there are N-1 restrictions implied by summability.

3. Symmetry. The matrix of compensated own- and cross-price substitution effects must be symmetric.

If the system of individual expenditure shares can be generated from an indirect utility function of the form (2.1), a necessary and sufficient condition for symmetry is that the

matrix B_{pp} must be symmetric. Without imposing this condition, we can write the individual expenditure shares in the form:

$$w_k = \frac{1}{D(p)} (\alpha_p + B_{pp} \ln \frac{p}{M_k} + B_{pA} A_k), \quad (k = 1, 2 \dots K).$$

Symmetry implies that the matrix of parameters B_{pp} must satisfy the restrictions:

$$B_{pp} = B'_{pp} \quad . \tag{2.9}$$

The total number of symmetry restrictions if $\frac{1}{2}N(N\text{-}1)$.

4. Nonnegativity. The individual expenditure shares must be nonnegative.

By summability the individual expenditure shares sum to unity, so that we can write:

$$w_k \geqq 0, \qquad (k = 1, 2 \dots K),$$

where $w_k \geq 0$ implies $w_{nk} \geq 0$, $(n = 1, 2 \dots N)$, and $w_k \neq 0$.

Since the translog indirect utility function is quadratic in the logarithms of prices, we can always choose the prices so that the individual expenditure shares violate the nonnegativity conditions. Accordingly, we cannot impose restrictions on the parameters of the translog indirect utility functions that would imply nonnegativity of the individual expenditure shares. Instead we consider restrictions on the parameters that imply monotonicity of the system of individual demand functions for all nonnegative expenditure shares.

5. Monotonicity. The matrix of compensated own- and cross-price substitution effects must be nonpositive definite.

We introduce the definition due to Martos [1969] of a **strictly merely positive subdefinite**

matrix, namely, a real symmetric matrix S such that:

$$xSx < 0$$

implies $Sx > 0$ or $Sx < 0$. A necessary and sufficient condition for monotonicity is either that the translog indirect utility function is homothetic or that B_{pp}^{-1} exists and is strictly merely positive subdefinite.[7]

To provide a basis for evaluating the impact of transfers among households on social welfare, we find it useful to represent household preferences by means of a utility function that is the same for all consuming units. For this purpose, we assume that the kth consuming unit maximizes its utility, say U_k, where:

$$U_k = U\left[\frac{x_{1k}}{m_1(A_k)}, \frac{x_{2k}}{m_2(A_k)} \cdots \frac{x_{Nk}}{m_N(A_k)}\right], \quad (k = 1, 2 \ldots K), \tag{2.10}$$

subject to the budget constraint:

$$M_k = \sum_{n=1}^{N} p_n x_{nk}, \qquad (k = 1, 2 \ldots K).$$

In this representation of consumer preferences the quantities $\{x_{nk}/m_n(A_k)\}$ can be regarded as **effective quantities consumed**, as proposed by Barten [1964]. The crucial assumption embodied in this representation is that differences in preferences among consumers enter the utility function U only through differences in the commodity specific household equivalence scales $\{m_n(A_k)\}$.[8]

Consumer equilibrium implies the existence of an indirect utility function, say V, that is the same for all consuming units. The level of utility for the kth consuming unit, say V_k, depends on the prices of individual commodities, the household equivalence scales,

and the level of total expenditure:

$$V_k = V\left[\frac{p_1\, m_1\,(A_k)}{M_k}, \frac{p_2\, m_2\,(A_k)}{M_k} \ldots \frac{p_N\, m_N\,(A_k)}{M_k}\right], \quad (k = 1, 2 \ldots K). \tag{2.11}$$

In this representation the prices $\{p_n m_n(A_k)\}$ can be regarded as **effective prices**. Differences in preferences among consuming units enter this indirect utility function only through the household equivalence scales $\{m_n(A_k)\}$ $(k = 1, 2 \ldots K)$.

To represent the translog indirect utility function (2.1) in terms of household equivalence scales, we require some additional notation:

$\ln \frac{p\, m(A_k)}{M_k}$ – vector of logarithms of ratios of effective prices

$\{p_n\, m_n(A_k)\}$ to total expenditure M_k of the kth consuming unit $(k = 1, 2 \ldots K)$.

$\ln m(A_k) = (\ln m_1(A_k), \ln m_2(A_k) \ldots \ln m_N(A_k))$ – vector of logarithms of the household equivalence scales of the kth consuming unit $(k = 1, 2 \ldots K)$.

We assume, as before, that the kth consuming unit allocates its expenditures in accord with the translog indirect utility function (2.1). However, we also assume that this function, expressed in terms of the effective prices $\{p_n\, m_n(A_k)\}$ and total expenditure M_k, is the same for all consuming units. The indirect utility function takes the form:

$$\ln V_k = \ln \frac{p\, m(A_k)}{M_k}' \alpha_p + \frac{1}{2} \ln \frac{p\, m(A_k)}{M_k}' B_{pp} \ln \frac{p\, m(A_k)}{M_k}, \quad (k = 1, 2 \ldots K). \tag{2.12}$$

Taking logarithms of the effective prices $\{p_n\, m_n(A_k)\}$, we can rewrite the indirect utility function (2.12) in the form:

$$\ln V_k = \ln m(A_k)'\alpha_p + \frac{1}{2}\ln m(A_k)'B_{pp}\ln m(A_k) + \ln \frac{p}{M_k}'\alpha_p \tag{2.13}$$

$$+\frac{1}{2}\ln \frac{p}{M_k}'B_{pp}\ln \frac{p}{M_k} + \ln \frac{p}{M_k}'B_{pp}\ln m(A_k), \qquad (k = 1, 2 \ldots K).$$

Comparing the representation (2.13) with the representation (2.6), we see that the term involving only the household equivalent scales must take the form:

$$F(A_k) = \ln m(A_k)'\,\alpha_p + \frac{1}{2}\ln m(A_k)'B_{pp}\ \ln m(A_k), \ (k = 1, 2 \ldots K). \tag{2.14}$$

Second, the term involving ratios of prices to total expenditure and the household equivalence scales must satisfy:

$$\ln \frac{p}{M_k}'B_{pA}\, A_k = \ln \frac{p}{M_k}'B_{pp}\ln m(A_k), \qquad (k = 1, 2 \ldots K). \tag{2.15}$$

for all prices and total expenditure.

The household equivalence scales $\{m_n(A_k)\}$ defined by (2.15) must satisfy the equation:

$$B_{pA}A_k = B_{pp}\ln m(A_k), \qquad (k = 1, 2 \ldots K). \tag{2.16}$$

Under monotonicity of the individual expenditure shares the matrix B_{pp} has an inverse, so that we can express the household equivalence scales in terms of the parameters of the translog indirect utility function – B_{pp} , B_{pA} – and the attributes $\{A_k\}$:

$$\ln m(A_k) = B_{pp}^{-1} B_{pA}\, A_k\,, \qquad (k = 1, 2 \ldots K). \tag{2.17}$$

We can refer to these scales as the **commodity specific translog household equivalence scales**.

Substituting the commodity specific equivalence scales (2.16) into the indirect utility function (2.13) we obtain a representation of the indirect utility function in terms of the attributes $\{A_k\}$:

$$\ln V_k = A'_k B'_{pA} B_{pp}^{-1} \alpha_p + \frac{1}{2} A'_k B'_{pA} B_{pp}^{-1} B_{pA} A_k \qquad (2.18)$$

$$+ \ln p' (\alpha_p + \frac{1}{2} B_{pp} \ln p + B_{pA} A_k) - D(p) \ln M_k , \quad (k = 1, 2 \ldots K).$$

This form of the translog indirect utility function is equivalent to the form (2.1) in that both generate the same system of individual demand functions. By requiring that the attributes A_k enter only through the commodity specific household equivalence scales, we have provided a specific form for the function $F(A_k)$ in (2.6).

Given the indirect utility function (2.18) for each consuming unit, we can express total expenditure as a function of prices, consumer attributes, and the level of utility:

$$\ln M_k = \frac{1}{D(p)} [A'_k B'_{pA} B_{pp}^{-1} \alpha_p + \frac{1}{2} A'_k B'_{pA} B_{pp}^{-1} B_{pA} A_k \qquad (2.19)$$

$$+ \ln p'(\alpha_p + \frac{1}{2} B_{pp} \ln p + B_{pA} A_k) - \ln V_k], \quad (k = 1, 2 \ldots K).$$

We can refer to this function as the **translog expenditure function**. The translog expenditure function gives the minimum expenditure required for the kth consuming unit to achieve the utility level V_k, given prices p (k = 1, 2 ... K).

We find it useful to introduce household equivalence scales that are not specific to a given commodity.[9] Following Deaton and Muellbauer [1980], we define a general household equivalence scale, say m_0, as follows:

$$m_0 = \frac{M_k [p\, m(A_k), V_k^0]}{M_0(p, V_k^0)}, \qquad (k = 1, 2 \ldots K). \qquad (2.20)$$

where M_k is the expenditure function for the kth household, M_0 is the expenditure function for a reference household with commodity specific equivalence scales equal to unity for all commodities, and p $m(A_k)$ is a vector of effective prices $\{p_n\ m_n(A_k)\}$.

The general household equivalence scale m_0 is the ratio between total expenditure required by the kth household and by the reference household for the same level of utility V_k^0 (k = 1, 2 ... K). This scale can be interpreted as the number of household equivalent members. The number of members depends on the attributes A_k of the consuming unit and on the prices p.

If each household has a translog indirect utility function, then the general household equivalence scale for the kth household takes the form:

$$\ln m_0 = \ln M_k - \ln M_0 , \tag{2.21}$$

$$= \frac{1}{D(p)} [\ln m\,(A_k)'\alpha_p + \frac{1}{2} \ln m\,(A_k)'B_{pp} \ln m\,(A_k) + \ln m\,(A_k)'B_{pp} \ln p],$$

$$(k = 1, 2 \dots K).$$

We can refer to this scale as the **general translog household equivalence scale**. The translog equivalence scale depends on the attributes A_k of the kth household and the prices p of all commodities, but is independent of the level of utility V_k^0.

Given the general translog equivalence scale, we can rewrite the indirect utility function (2.18) in the form:

$$\ln V_k = \ln p'\alpha_p + \frac{1}{2} \ln p'B_{pp} \ln p - D(p) \ln [M_k/m_0(p, A_k)], \ (k = 1, 2 \dots K). \tag{2.22}$$

The level of utility for the kth consuming unit depends on prices p and total expenditure per household equivalent member $M_k/m_0(p, A_k)$ (k = 1, 2 ... K). Similarly, we can rewrite

the expenditure function (2.19) in the form:

$$\ln M_k = \frac{1}{D(p)} [\ln p'(\alpha_p + \frac{1}{2} B_{pp} \ln p) - \ln V_k] + \ln m_0(p, A_k),$$

$$(k = 1, 2 \dots K). \tag{2.23}$$

Total expenditure required by the kth consuming unit to attain the level of utility V_k depends on prices p and the number of household equivalent members $m_0(p, A_k)$ $(k = 1, 2 \dots K)$.

To construct an econometric model of aggregate consumer behavior based on exact aggregation we obtain aggregate expenditure shares, say w, by multiplying individual expenditure shares (2.5) by expenditure for each consuming unit, adding over all consuming units, and dividing by aggregate expenditure, $M = \sum_{k=1}^{K} M_k$:

$$w = \frac{\Sigma M_k w_k}{M}. \tag{2.24}$$

The aggregate expenditure shares can be written:

$$w = \frac{1}{D(p)} (\alpha_p + B_{pp} \ln p - B_{pp} i \frac{\Sigma M_k \ln M_k}{M} + B_{pA} \frac{\Sigma M_k A_k}{M}). \tag{2.25}$$

The aggregate expenditure patterns depend on the distribution of expenditure over all consuming units through summary statistics of the joint distribution of expenditures and attributes – $\Sigma M_k \ln M_k / M$ and $\{\Sigma M_k A_k / M\}$. Systems of individual expenditure shares (2.5) for consuming units with identical demographic characteristics can be recovered in one and only one way from the system of aggregate expenditure shares (2.25).

To summarize: Systems of individual expenditure shares (2.5) can be recovered in one and only one way from the system of aggregate expenditure shares (2.25). Given a system

of individual exenditure shares (2.5) that is integrable, we can recover the indirect utility function (2.22). This indirect utility function provides a cardinal measure of utility. We obtain measures of utility for all consuming units by deriving indirect utility functions from the fitted systems of individual expenditure shares.

3. Econometric Model

In this section we present the empirical results of implementing the econometric model of consumer behavior described in Section 2. We divide consumer expenditures among five commodity groups:

1. Energy: Expenditures on electricity, natural gas, heating oil and gasoline.

2. Food: Expenditures on all food products, including tobacco and alcohol.

3. Consumer Goods: Expenditures on all other nondurable goods included in consumer expenditures.

4. Capital Services: The service flow from consumer durables and the service flow from housing.

5. Consumer Services: Expenditures on consumer services, such as car repairs, medical services, entertainment, and so on.

We employ the following demographic characteristics as attributes of individual households:

1. Family size: 1, 2, 3, 4, 5, 6 and 7 or more persons.

2. Age of head: 16-24, 25-34, 35-44, 45-54, 55-64, 65 and over.

3. Region of residence: Northeast, North Central, South and West.

4. Race: White, nonwhite.

5. Type of residence: Urban, rural.

Our cross-section observations on individual expenditures for each commodity group and on demographic characteristics of individual households are for the year 1972 from the 1972-1973 Survey of Consumer Expenditures (CES).[10] Our time series observations are based on data on personal consumption expenditures from the United States National Income and Product Accounts (NIPA) for the years 1958 to 1974.[11] Prices for each commodity group are defined in terms of translog price indexes computed from detailed prices

included in NIPA for each year. We employ time series data on the distribution of expenditures over all households and among demographic groups based on **Current Population Reports.**[12]

In our application we treat the expenditure shares for five commodity groups as endogenous variables, so that we estimate four equations. As unknown parameters we have four elements of the vector α_p, four expenditure coefficients of the vector $B_{pp}i$, 16 attribute coefficients for each of the four equations in the matrix B_{pA}, and 10 price coefficients in the matrix B_{pp}, which is constrained to be symmetric. The expenditure coefficients are sums of price coefficients in the corresponding equation, so that we have a total of 82 unknown parameters. We estimate the complete model, subject to inequality restrictions implied by monotonicity of the individual expenditure shares, by pooling time series and cross-section data.[13] The results are given in Table 1.

The impacts of changes in total expenditures and in demographic characteristics of the individual household are estimated very precisely. This reflects the fact that estimates of the expenditure and demographic effects incorporate a relatively large number of cross-section observations. The impacts of prices enter through the denominator of the equations for expenditure shares; these price coefficients are estimated very precisely since they also incorporate cross-section data. Finally, the price impacts also enter through the numerators of equations for the expenditure shares. These parameters are estimated somewhat less precisely, since they are based on a much smaller number of time series observations on prices.

To summarize: We have implemented an econometric model of aggregate consumer behavior by combining time series and cross-section data for the United States. This model allocates personal consumption expenditures among five commodity groups -- energy, food, other consumer goods, capital services, and other consumer services. Households are classified by five sets of demographic characteristics -- family size, age of head, region of residence, race, and urban versus rural residence.

TABLE 1. Pooled Estimation Results

Notation:

CONST	=	constant term.
1n PEN	=	coefficient of log of price of energy.
1n PF	=	coefficient of log of price of food.
1n PCG	=	coefficient of log of price of consumer goods.
1n PK	=	coefficient of log of price of capital services.
1n PCS	=	coefficient of log of price of consumer services.
1n M	=	coefficient of log of total expenditure.
S2	=	coefficient of dummy for family of size 2.
S3	=	coefficient of dummy for family of size 3.
S4	=	coefficient of dummy for family of size 4.
S5	=	coefficient of dummy for family of size 5.
S6	=	coefficient of dummy for family of size 6.
S7+	=	coefficient of dummy for family of size 7 or more.
A25-34	=	coefficient of dummy for age between 25 and 34.
A35-44	=	coefficient of dummy for age between 35 and 44.
A45-54	=	coefficient of dummy for age between 45 and 54.
A55-64	=	coefficient of dummy for age between 55 and 64.
A65+	=	coefficient of dummy for age 65 and over.
RNC	=	coefficient of dummy for family living in North Central.
RS	=	coefficient of dummy for family living in South.
RW	=	coefficient of dummy for family living in West.
NW	=	coefficient of dummy for nonwhite family.
RUR	=	coefficient of dummy for family living in rural area.

$$
\begin{aligned}
D(p) = -1 &- \underset{(.000997)}{.03491} \ln PEN - \underset{(.00238)}{.08171} \ln PF + \underset{(.00214)}{.06189} \ln PCG \\
&- \underset{(.00300)}{.002060} \ln PK + \underset{(.00233)}{.05679} \ln PCS
\end{aligned}
$$

TABLE 1. Energy – Continued

Parameter	Estimate	Standard Error
CONST	–.3754	.00923
ln PEN	.09151	.0134
ln PF	–.1441	.0214
ln PCG	–.06455	.0127
ln PK	.07922	.0171
ln PCS	.003061	.0138
ln M	.03491	.000997
S2	–.02402	.00139
S3	–.02971	.00163
S4	–.03144	.00178
S5	–.03255	.00206
S6	–.03606	.00249
S7+	–.02977	.00266
A25-34	.0002010	.00197
A35-44	–.006703	.00210
A45-54	–.01155	.00199
A55-64	–.01372	.00199
A65+	–.005487	.00196
RNC	–.003277	.00131
RS	.0001280	.00131
RW	.01281	.00140
NW	.01300	.00170
RUR	–.03057	.00134

TABLE 1. Food – Continued

Parameter	Estimate	Standard Error
CONST	–.8917	.0215
1n PEN	–.1441	.0214
1n PF	.3118	.0428
1n PCG	.05547	.0215
1n PK	–.1982	.0334
1n PCS	–.1066	.0259
1n M	.08171	.00238
S2	–.04859	.00333
S3	–.06730	.00390
S4	–.08881	.00428
S5	–.1108	.00496
S6	–.1185	.00598
S7+	–.1471	.00639
A25-34	–.04393	.00474
A35-44	–.08221	.00504
A45-54	–.09604	.00478
A55-64	–.1034	.00477
A65+	–.08833	.00470
RNC	.08173	.00315
RS	.01213	.00314
RW	.01856	.00337
NW	.006274	.00409
RUR	–.001793	.00323

TABLE 1. Consumer Goods – Continued

Parameter	Estimate	Standard Error
CONST	.4053	.0194
1n PEN	–.06455	.0127
1n PF	.05547	.0215
1n PCG	.2301	.0269
1n PK	–.1056	.0195
1n PCS	–.05354	.0271
1n M	–.06189	.00214
S2	–.005594	.00300
S3	–.006290	.00351
S4	–.001941	.00385
S5	.004522	.00446
S6	.01059	.00539
S7+	.01495	.00575
A25-34	–.02311	.00426
A35-44	–.01916	.00454
A45-54	–.005279	.00431
A55-64	–.009068	.00429
A65+	–.01722	.00423
RNC	–.02098	.00283
RS	–.03553	.00283
RW	–.009928	.00304
NW	–.02648	.00368
RUR	–.01122	.00290

TABLE 1. Capital Services – Continued

Parameter	Estimate	Standard Error
CONST	–.4658	.0270
1n PEN	.07922	.0171
1n PF	–.1982	.0334
1n PCG	–.1056	.0195
1n PK	.2038	.0368
1n PCS	.01869	.0165
1n M	.002060	.00300
S2	.07355	.00421
S3	.09982	.00493
S4	.1148	.00541
S5	.1253	.00626
S6	.1284	.00756
S7+	.1369	.00807
A25-34	.04362	.00599
A35-44	.08503	.00637
A45-54	.1166	.00605
A55-64	.1395	.00603
A65+	.1296	.00595
RNC	.02767	.00398
RS	.05528	.00397
RW	–.004132	.00427
NW	–.003539	.00517
RUR	.05588	.00408

TABLE 1. Consumer Services – Concluded

Parameter	Estimate	Standard Error
CONST	.3277	.0211
ln PEN	.003061	.0138
ln PF	–.1066	.0259
ln PCG	–.05354	.0271
ln PK	.01869	.0165
ln PCS	.1952	.0375
ln M	–.05679	.00233
S2	.004666	.00327
S3	.003483	.00383
S4	.007338	.00420
S5	.01357	.00486
S6	.01561	.00587
S7+	.02508	.00627
A25-34	.02321	.00465
A35-44	.02304	.00495
A45-54	–.003805	.00470
A55-64	–.01332	.00468
A65+	–.01863	.00462
RNC	–.02214	.00309
RS	–.03200	.00308
RW	–.01731	.00331
NW	.01074	.00401
RUR	–.01229	.00317

Convergence after 3 iterations.

SSR = 37387.12.

Convergence criterion = .00001.

4. Individual Cost-of-Living Indexes

In this section we outline a methodology for translating changes in prices into measures of change in the individual cost of living. The first step in measuring the cost of living for an individual consuming unit is to select a representation for the individual welfare function. We assume that individual welfare for the kth consuming unit, say W_k $(k = 1, 2 \dots K)$, is equal to the logarithm of the translog indirect utility function (2.22):

$$W_k = \ln V_k , \tag{4.1}$$

$$= \ln p'\alpha_p + \frac{1}{2}\ln p' B_{pp} \ln p - D(p) \ln [M_k / m_0(p, A_k)], \quad (k = 1, 2 \dots K).$$

Following Konüs [1939], we define the cost-of-living index for the kth consuming unit, say P_k $(k = 1, 2 \dots K)$, as the ratio of two levels of total expenditure:

$$P_k(p^1, p^0, W_k^0, A_k) = \frac{M_k(p^1, W_k^0, A_k)}{M_k(p^0, W_k^0, A_k)}, \quad (k = 1, 2 \dots K). \tag{4.2}$$

In this ratio the numerator is the expenditure required to attain the base level of individual welfare W_k^0 at the current price system p^1; the denominator is the base level of expenditure M_k^0.

The translog indirect utility function (2.22) and the translog expenditure function (2.23) can be employed in implementing the individual cost-of-living index (4.2).[14] First, we can express the base level of individual welfare W_k^0 in terms of the translog indirect utility function:

$$W_k^0 = \ln p^{0\prime} \alpha_p + \frac{1}{2}\ln p^{0\prime} B_{pp} \ln p^0 - D(p^0) \ln [M_k^0 / m_0(p^0, A_k)], \quad (k = 1, 2 \dots K). \tag{4.3}$$

In this expression $m_0(p^0, A_k)$ is the base level of the general translog household equivalence scale (2.21) and M_k^0 is the base level of total expenditure.

Using the translog expenditure function, we can express the individual cost-of-living index (4.2) in the form:

$$\ln P_k(p^1, p^0, W_k^0, A_k) = \frac{1}{D(p^1)} [\ln p^{1'} (\alpha_p + \frac{1}{2} B_{pp} \ln p^1) - W_k^0]$$
$$+ \ln m_0(p^1, A_k) - \ln M_k^0 , \quad (k = 1, 2 \dots K). \tag{4.4}$$

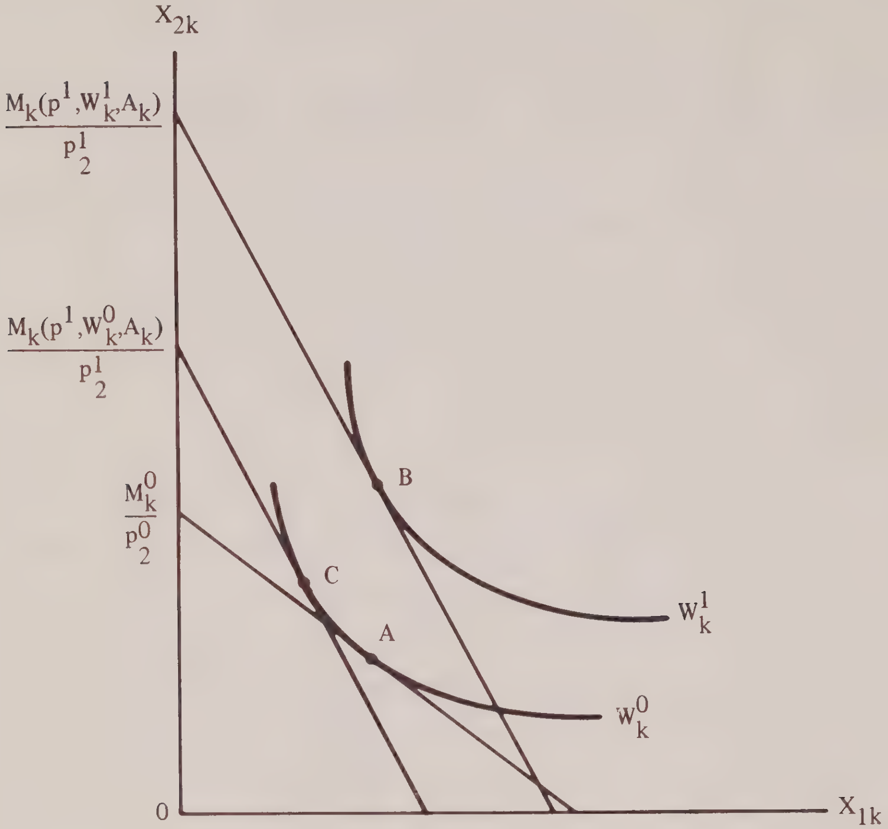
We can refer to the index P_k as the **translog individual cost-of-living index**. If the translog index is greater than unity and total expenditure is constant, then the welfare of the consuming unit is decreased by the change in prices.

We can illustrate the measurement of the individual cost-of-living index by representing the impact of a change in prices in diagrammatic form. For simplicity we consider the case of two commodities ($N = 2$). In Diagram 1 we have depicted the indifference map of the kth household. Consumer equilibrium at the base price system p^0 is represented by the point A with base level of individual welfare W_k^0. The corresponding level of total expenditure M_k^0, divided by the base price of the second commodity p_2^0, is given on the vertical axis. This axis provides a representation of total expenditure in terms of units of the second commodity.

Consumer equilibrium after the change in prices is represented by the point B with associated level of individual welfare W_k^1. The level of total expenditure associated with the change in prices M_k^1, divided by the current price of the second commodity p_2^1, is given, as before, on the vertical axis. Finally, the level of total expenditure required to attain the base level of individual welfare W_k^0 at current prices $M(p^1, W_k^0, A_k)$ corresponds to consumer equilibrium at the point C. The individual cost- of-living index is given by ratio of the distances on the vertical axis corresponding to the consumer equilibrium at points C and A, multiplied by the ratio of prices of the second commodity at the two points p_2^1 / p_2^0.

As a further illustration of the individual cost-of-living index, we analyze the changes in prices over the period 1958-1978, using prices for 1972 as the base price system. For this purpose we employ the econometric model of aggregate consumer behavior presented

Diagram 1. Individual Cost of Living Index.



in Section 3. This model is based on time series and cross-section data on personal consumption expenditures for the United States, broken down by five commodity groups. The five commodity groups are energy, food, other consumer goods, capital services and other consumer services.

Using the translog individual cost-of-living index (4.4), we can assess the impact of price changes on households with different base levels of individual welfare and different demographic characteristics. For this purpose we set the base level of individual welfare at the levels attained in 1972 with half mean expenditure in that year of \$4,467, mean expenditure or \$8,934, and twice mean expenditure or \$17,868. We present translog individual cost-of-living indexes for the period 1958-1978 for white and nonwhite households with urban and rural residences in Appendix Table 1. Within each of these groups we consider

TABLE 2. Changes in Individual Cost-of-Living Indexes (annual percentage rates)

Year	Urban		Rural	
	White	Nonwhite	White	Nonwhite
1959	−0.02	−0.05	0.24	0.21
1960	2.94	2.95	2.67	2.68
1961	0.84	0.83	0.86	0.85
1962	1.03	1.02	1.01	1.00
1963	0.46	0.46	0.62	0.62
1964	2.19	2.23	1.91	1.95
1965	3.26	3.26	2.94	2.93
1966	2.97	2.95	3.01	2.99
1967	0.73	0.74	1.09	1.10
1968	3.16	3.21	3.24	3.29
1969	6.18	6.23	5.76	5.80
1970	1.48	1.43	2.15	2.10
1971	3.78	3.76	3.77	3.75
1972	7.05	7.07	6.29	6.31
1973	8.21	8.14	8.00	7.93
1974	8.87	8.56	10.33	10.03
1975	4.65	4.60	5.40	5.34
1976	6.39	6.43	6.10	6.14
1977	8.36	8.31	7.91	7.86
1978	7.04	7.02	6.90	6.88

Region = Northeast.
Size = 5.
Age = 35-44.
Expenditure = \$4467. in 1972.

TABLE 2. Changes in Individual Cost-of-Living Indexes (annual percentage rates) – Continued

Year	Urban		Rural	
	White	Nonwhite	White	Nonwhite
1959	0.08	0.06	0.35	0.33
1960	2.97	2.97	2.69	2.70
1961	0.84	0.83	0.86	0.86
1962	1.06	1.05	1.04	1.03
1963	0.49	0.49	0.65	0.65
1964	2.21	2.24	1.93	1.96
1965	3.18	3.17	2.85	2.84
1966	2.86	2.84	2.90	2.88
1967	0.86	0.87	1.23	1.24
1968	3.26	3.31	3.35	3.39
1969	6.21	6.25	5.79	5.83
1970	1.49	1.44	2.16	2.10
1971	3.91	3.89	3.90	3.88
1972	6.96	6.97	6.20	6.22
1973	7.71	7.63	7.50	7.43
1974	8.15	7.84	9.61	9.31
1975	4.68	4.63	5.43	5.37
1976	6.79	6.82	6.50	6.53
1977	8.48	8.43	8.03	7.98
1978	6.70	6.68	6.55	6.53

Region = Northeast.
Size = 5.
Age = 35-44.
Expenditure = $8,934 in 1972.

TABLE 2. Changes in Individual Cost-of-Living Indexes (annual percentage rates) – Concluded

Year	Urban		Rural	
	White	Nonwhite	White	Nonwhite
1959	0.20	0.17	0.47	0.44
1960	2.99	2.99	2.72	2.72
1961	0.84	0.84	0.87	0.86
1962	1.08	1.07	1.07	1.06
1963	0.52	0.52	0.68	0.68
1964	2.23	2.26	1.95	1.98
1965	3.09	3.08	2.76	2.75
1966	2.75	2.73	2.79	2.78
1967	1.00	1.01	1.37	1.38
1968	3.37	3.42	3.45	3.50
1969	6.24	6.28	5.81	5.85
1970	1.50	1.44	2.16	2.11
1971	4.04	4.02	4.03	4.01
1972	6.86	6.88	6.11	6.12
1973	7.21	7.13	7.00	6.92
1974	7.43	7.12	8.90	8.59
1975	4.71	4.66	5.45	5.40
1976	7.18	7.21	6.89	6.92
1977	8.59	8.55	8.15	8.10
1978	6.35	6.33	6.21	6.19

Region = Northeast.
Size = 5.
Age = 35-44.
Expenditure = $17,868 in 1972.

families of size five with a head of household aged 35-44, living in the Northeast region of the United States.

We present rates of inflation calculated from the translog individual cost-of-living indexes in Table 2. For example, the change in the logarithm of the translog index between 1958 and 1959 for a white, urban family with total expenditure equal to $4,467 in 1972 is -0.02 percent. The corresponding change for a white, rural family is 0.24 percent. For nonwhite households the change in cost of living for urban households is -0.05 percent, while the change for rural households is 0.21 percent. If we compare translog individual cost-of-living indexes for households with different base levels of total expenditure, we find that rates of inflation are greater for higher levels of expenditure for the periods 1958-1964, 1966-1971, and 1974-1977. For the remainder of the period -- 1964-1966, 1971-1974, and 1977-78 -- rates of inflation are greater for lower levels of expenditure.

We can compare rates of inflation for households with different demographic characteristics. For example, a white, urban family with mean base level of expenditure experienced a change of 8.15 percent in the cost-of-living between 1973 and 1974. By comparison a white, rural household at the same level of expenditure had a change of 9.61 percent over the same period. On the basis of the results presented in Table 2 we conclude that there are substantial differences in rates of inflation for households with different base levels of individual welfare and with different demographic characteristics.

To summarize: We have defined the individual cost of living index as the ratio of the total expenditure required to attain a base level of individual welfare at current prices to the base level of expenditure. Using the translog indirect utility function (2.22) and the translog expenditure function (2.23), we implement this definition by means of the translog individual cost of living index (4.4). We find that changes in the individual cost-of-living vary substantially for households with different base levels of welfare and different demographic characteristics over the period 1958-1978.

5. Social Welfare Functions

Our next objective is to generate a class of possible social welfare functions that can

provide the basis for social cost of living measurement. For this purpose we must choose social welfare functions capable of expressing the implications of a variety of different ethical judgements. To facilitate comparisons with alternative approaches, we employ the axiomatic framework for social choice used by Arrow [1963], Sen [1970], and Roberts [1980a] in proving the impossibility of a nondictatorial social ordering.

We consider the set of all possible social orderings over the set of social states, say X, and the set of all possible real-valued individual welfare functions, say W_k $(k = 1, 2 \ldots K)$. A social ordering, say R, is a complete, reflexive, and transitive ordering of social states. A social state is described by the quantities consumed of N commodity groups by K individuals. The individual welfare function for the kth individual W_k $(k = 1, 2 \ldots K)$ is defined on the set of social states X and gives the level of individual welfare for that individual in each state.

To describe social orderings in greater detail we find it useful to introduce the following notation:

x_{nk} – the quantity of the nth commodity group consumed by the kth consuming unit $(n = 1, 2 \ldots N; k = 1, 2 \ldots K)$.

x – a matrix with elements $\{x_{nk}\}$ describing the social state.

$u = (W_1, W_2 \ldots W_K)$ – a vector of individual welfare functions of all K individuals.

Following Sen [1970, 1977] and Hammond [1976] we define a **social welfare functional,** say f, as a mapping from the set of individual welfare functions to the set of social orderings, such that $f(u') = f(u)$ implies $R' = R$, where:

$$u = [W_1(x), W_2(x) \ldots W_K(x)],$$

$$u' = [W_1'(x), W_2'(x) \ldots W_K'(x)],$$

for all $x \in X$. Similarly, we define L_k ($k = 1, 2 \ldots K$) as the **set of admissible individual welfare functions** for the kth individual and L as the Cartesian product $\prod_{k=1}^{K} L_k$. Finally let $\underline{L}$ be the partition of L such that all elements of $\underline{L}$ yield the same social ordering.

We can describe a social ordering in terms of the following properties of a social welfare functional:

1. Unrestricted Domain. The social welfare functional f is defined for all possible vectors of individual welfare functions u.

2. Independence of Irrelevant Alternatives. For any subset A contained in X, if $u(x) = u'(x)$ for all $x \in A$, then $R{:}A = R'{:}A$, where $R = f(u)$ and $R' = f(u')$ and $R{:}A$ is the social ordering over the subset A.

3. Positive Association. For any vectors of individual welfare functions u and u′, if for all y in X-x, such that:

$$W'_k(y) = W_k(y),$$

$$W'_k(x) > W_k(x), \qquad (k = 1, 2 \ldots K),$$

then xPy implies xP′y and yP′x implies yPx, where P is a strict ordering of social states.

4. Nonimposition. For all x, y in X there exist u, u′ such that xPy and yP′x.

5. Cardinal Full Comparability. The set of admissible individual welfare functions that yield the same social ordering $\underline{L}$ is defined by:

$$\underline{L} = \left\{u' : W'_k(x) = \alpha + \beta W_k(x),\ \beta > 0,\ k = 1, 2 \ldots K\right\},$$

and $f(u') = f(u)$ for all $u' \in \underline{L}$.

Cardinal full comparability implies that social orderings are invariant with respect to any positive affine transformation of the individual welfare functions $\{W_k\}$ that is the same for all individuals. By contrast Arrow requires ordinal noncomparability,[15] which

implies that social orderings are invariant with respect to monotone increasing transformations of the individual welfare functions that may differ among individuals:

5′. Ordinal Noncomparability. The set of individual welfare functions that yield the same social ordering $\underline{L}$ is defined by:

$$\underline{L} = \left\{u' : W'_k(x) = \phi_k\,[W_k(x)],\ \phi_k \text{ increasing},\ k = 1, 2 \ldots K\right\},$$

and $f(u') = f(u)$ for all u' in $\underline{L}$.

The properties of a social welfare functional corresponding to unrestricted domain and independence of irrelevant alternatives are used by Arrow in proving the impossibility of a nondictatorial social ordering:

4′. Nondictatorship. There is no individual k such that for all $x, y \in X$, $W_k(x) > W_k(y)$ implies xPy.

Under ordinal noncomparability the assumptions of positive association and nonimposition employed by Arrow imply the weak Pareto principle:

3′Pareto Principle. For any $x, y \in X$, if $W_k(x) > W_k(y)$ for all individuals $(k = 1, 2 \ldots K)$, then xPy.

If a social welfare functional f has the properties of unrestricted domain, independence of irrelevant alternatives, the weak Pareto principle, and ordinal noncomparability, then no nondictatorial social ordering is possible. This result is Arrow's impossibility theorem. Since it is obvious that the class of dictatorial social orderings is too narrow to provide an adequate basis for expressing the implications of alternative ethical judgments, we propose to generate a class of social welfare functions suitable for the evaluation of alternative economic policies by weakening Arrow's assumptions.

We first consider weakening the assumption of ordinal noncomparability of individual welfare functions. Sen [1970] has shown that Arrow's conclusion that no nondictatorial social ordering is possible is preserved by replacing ordinal noncomparability by cardinal

noncomparability. This implies that social orderings are invariant with respect to positive affine transformations of the individual welfare functions that may differ among individuals:

5′′. Cardinal Noncomparability. The set of individual welfare functions that yield same social ordering $\underline{L}$ is defined by:

$$\underline{L} = \left\{u': W_k'(x) = \alpha_k + \beta_k W_k(x), \beta_k > 0, k = 1, 2 \ldots K\right\},$$

and $f(u') = f(u)$ for all u' in $\underline{L}$.

However, d'Aspremont and Gevers [1977], Deschamps and Gevers [1978], Maskin [1978] and Roberts [1980b] have shown that we obtain an interesting class of nondictatorial social orderings by requiring cardinal unit comparability of individual welfare functions, which implies that social orderings are invariant with respect to positive affine transformations with units that are the same for all individuals:

5′′′. Cardinal Unit Comparability. The set of individual welfare functions that yield the same social ordering $\underline{L}$ is defined by:

$$\underline{L} = \left\{u': W_k'(x) = \alpha_k + \beta W_k(x), \beta > 0, k = 1, 2 \ldots K\right\},$$

and $f(u') = f(u)$ for all u' in $\underline{L}$.

If a social welfare functional f has the properties of unrestricted domain, independence of irrelevant alternatives, the weak Pareto principle, and cardinal unit comparability, there exist social orderings and a continuous real-valued social welfare function, say W, such that if $W[u(x)] > W[u(y)]$, then xPy. Furthermore, the social welfare function can be represented in the form:

$$W[u(x)] = \sum_{k=1}^{K} a_k W_k(x). \tag{5.1}$$

If we add the assumption that the social welfare function has the property of anonymity, that is, no individual is given greater weight than any other individual in determining the level of social welfare, then the social welfare function W in (5.1) must be symmetric in the individual welfare functions $\{W_k\}$. The property of anonymity incorporates a notion of horizontal equity into the representation of social orderings.

Under anonymity the function W in (5.1) reduces to the sum of individual welfare functions and takes the form of a utilitarian social welfare function. Utilitarian social welfare functions have been employed extensively in applications of welfare economics, especially in the measurement of inequality by methods originated by Atkinson [1970] and Kolm [1969, 1976a, 1976b], in the design of optimal income tax schedules along the lines pioneered by Mirrlees [1971], and in the evaluation of alternative economic policies by Arrow and Kalt [1979].

The approach to the measurement of social welfare based on a utilitarian social welfare function provides a worthwhile starting point for applications. Harsanyi [1976] and Ng [1975] have pointed out that distributional considerations can be incorporated into a utilitarian social welfare function through the representation of individual welfare functions. However, Sen [1973, p.18] has argued that a utilitarian social welfare function does not take appropriate account of the distribution of welfare among individuals:

> The distribution of welfare between persons is a relevant aspect of any problem of income distribution, and our evaluation of inequality will obviously depend on whether we are concerned only with the loss of the sum of individual utilities through a bad distribution of income, or also with the inequality of welfare levels of different individuals.

To broaden the range of possible social orderings we can require cardinal full comparability of individual welfare functions, as defined above. Roberts [1980b] has shown that a social welfare functional f with the properties of unrestricted domain, independence of irrelevant alternatives, the weak Pareto principle, and cardinal full comparability implies

the existence of a social welfare function that takes the form:

$$W[u(x)] = \bar{W}(x) + g[u(x) - \bar{W}(x)\, i], \tag{5.2}$$

where i is a vector of ones, the function $\bar{W}(x)$ corresponds to average individual welfare:

$$\bar{W}(x) = \sum_{k=1}^{K} a_k W_k(x),$$

and g(x) is a linear homogeneous function of deviations of levels of individual welfare from the average.[16]

If the function g(x) in the representation (5.2) of the social welfare function is identically equal to zero, then the social welfare function reduces to the form (5.1). If the function g(x) is not identically zero, then the social welfare function incorporates both a measure of average individual welfare and a measure of inequality in the distribution of individual welfare. We conclude that the class of possible social welfare functions (5.2) includes utilitarian welfare functions, but also includes functions that are not subject to the objections that can be made to utilitarianism.

Although Roberts [1980b] has succeeded in broadening the class of possible social welfare functions beyond those consistent with utilitarianism, the social welfare functions (5.2) are subject to an objection raised by Sen [1973].[17] Information about alternative social states enters only through the individual welfare functions $\{W_k\}$. Sen refers to this property of a social welfare function f as **welfarism**. Welfarism rules out characteristics of a social state that are conceivably relevant for social orderings, but that cannot be incorporated into the social welfare function through the individual welfare functions.

Roberts [1980b] has suggested the possibility of further weakening of Arrow's assumptions in order to incorporate nonwelfare characteristics of social states.[18] For this purpose we can replace the weak Pareto principle by positive association and nonimposition, as defined above. We retain the assumptions of unrestricted domain, independence of irrelevant alternatives, and cardinal full comparability of measures of individual welfare. We can partition the set of social states X into subsets, such that all states within each subset

have the same nonwelfare characteristics. For each subset there exists a social ordering that can be represented by a social welfare function of the form (5.2).

Under the assumptions we have outlined there exists a social ordering for the set of all social states that can be represented by a social welfare function of the form:

$$W(u, x) = F\left\{\bar{W}(x) + g[x, u(x) - \bar{W}(x)\, i], x\right\}, \qquad (5.3)$$

where the function $\bar{W}(x)$ corresponds to average individual welfare:

$$\bar{W}(x) = \sum_{k=1}^{K} a_k(x)\, W_k(x).$$

As before, the function g is a linear homogeneous function of deviations of levels of individual welfare from average welfare.

The class of social welfare functions (5.3) incorporates nonwelfare characteristics of social states through the weights $\{a_k(x)\}$ in average individual welfare $\bar{W}(x)$, through the function $g(x)$, which depends directly on the social state x as well as on deviations of levels of individual welfare from the average welfare, and through the function F, which depends directly on the social state x and on the sum of the functions $\bar{W}(x)$ and $g(x)$. This class includes social welfare functions that are not subject to the objections that can be made to welfarism.

At this point we have generated a class of possible social welfare functions capable of expressing the implications of a variety of different ethical judgments. In order to choose a specific social welfare function, we must narrow the range of possible ethical judgments

by imposing further requirements on the class of possible social welfare functions. First, we must limit the dependence of the function F(x) in (5.3) on the characteristics of alternative social states. Second, we must select a form for the function g(x) in (5.3), which depends on deviations of levels of individual welfare from average welfare $\bar{W}(x)$. Finally, we must choose representations of the individual welfare functions $\{W_k(x)\}$ that provide cardinal full comparability.

We first rule out the dependence of the function F(x) in (5.3) on characteristics of social states that do not enter through the functions $\bar{W}(x)$ and g(x). This restriction reduces F to a function of a single variable $\bar{W} + g$. We obtain an ordinal measure of social welfare by permitting the function F to be any monotone increasing transformation. To obtain a cardinal measure of social welfare we observe that the function $\bar{W}(x) + g$ is homogeneous of degree one in the individual welfare functions $\{W_k(x)\}$. All representations of the social welfare function that preserve this property can be written in the form:

$$W(u, x) = \beta[\bar{W}(x) + g(x)], \beta > 0. \tag{5.4}$$

We conclude that only positive, homogeneous, affine transformations are permitted.

The restrictions embodied in the class of social welfare functions (5.4) do not reduce social welfare to a function of the individual welfare functions $\{W_k(x)\}$ alone, since the weights $\{a_k(x)\}$ in average individual welfare $\bar{W}(x)$ and the function g(x) depend on nonwelfare characteristics of the social state x. However, these social welfare functions are homogeneous of degree one in levels of individual welfare. This implies that doubling the welfare of each individual will double social welfare, holding nonwelfare characteristics of the social state constant. Blackorby and Donaldson [1982] refer to this class of social welfare functions as **distributionally homothetic**.[19]

We impose a second set of requirements on the class of social welfare functions (5.3) by selecting an appropriate form for the function g(x). In particular, we require that this function is additive in deviations of individual welfare functions $\{W_k(x)\}$ from average welfare $\bar{W}(x)$. Since the function g(x) is homogeneous of degree one, it must be a mean

value function of order $\rho(x)$:[20]

$$g[x, u(x) - \bar{W}(x)\, i] = -\gamma(x)\Big[\sum_{k=1}^{K} b_k(x)\, |W_k - \bar{W}|^{-\rho(x)}\Big]^{-\frac{1}{\rho(x)}}, \qquad (5.5)$$

where:

$$\gamma(x) > 0,\ \rho(x) \leq -1,\ \sum_{k=1}^{K} b_k(x) = 1,\ 0 < b_k(x) < 1,\ (k = 1, 2 \dots K).$$

Under these restrictions the function g(x) is negative, except at the point of perfect equality $W_k = \bar{W}$ $(k = 1, 2 \dots K)$, where it is zero.

The function (x) in the representation (5.5) determines the curvature of the social welfare function in the individual welfare functions $\{W_k(x)\}$. We can refer to this function as the **degree of aversion to inequality**. We assume that this function is constant, so that the corresponding social welfare function W(u,x) is characterized by a constant degree of aversion to inequality. To complete the selection of an appropriate form for the social welfare function we must choose appropriate weights $\{a_k(x)\}$ for average individual welfare $\bar{W}(x)$ and $\{b_k(x)\}$ for the measure of equality g(x). We find it natural to require that the two sets of weights are the same.

To incorporate a notion of horizontal equity into the social welfare functions (5.5) we can impose a weak form of the property of anonymity. In particular, we require that no individual is given greater weight in the social welfare function than any other individual with an identical individual welfare function. This implies that the social welfare function is symmetric in the levels of individual welfare for identical individuals. The weights $\{a_k(x)\}$ in average welfare $\bar{W}(x)$ and the measure of equality g(x) must be the same for identical individuals.

Under the restrictions presented up to this point the social welfare function W takes

the form:

$$W(u,x) = \bar{W} - \gamma(x) \left[\sum_{k=1}^{K} a_k(x) | W_k - \bar{W} |^{-\rho} \right]^{-\frac{1}{\rho}} \tag{5.6}$$

where:

$$\bar{W}(x) = \sum_{k=1}^{K} a_k(x) \, W_k(x).$$

The condition of positive association requires that an increase in all levels of individual welfare must increase social welfare. This condition implies that the average level of individual welfare $\bar{W}$ must increase by more than the function g(x), whatever the initial distribution of individual welfare. We assume that the function $\gamma(x)$ in (5.6) must take the maximum value consistent with positive association, so that:

$$\gamma(x) = \left\{ 1 + \left[\frac{\sum_{k=1}^{K} a_k(x)}{a_j(x)} \right]^{-(\rho+1)} \right\}^{\frac{1}{\rho}}, \tag{5.7}$$

where:

$$a_j(x) = \min_k a_k(x), \qquad (k = 1, 2 \dots K),$$

for the social state x.

To complete the selection of a social welfare function W(u,x) we require that the individual welfare functions $\{W_k\}$ in (5.3) must be invariant with respect to any positive affine transformation that is the same for all households.[21] Under this assumption the logarithm of the translog indirect utility function is a cardinal measure of individual welfare

with full comparability among households. The social welfare function takes the form:

$$W(u,x) = \ln \bar{V} - \gamma(x) \left[\sum_{k=1}^{K} a_k(x) |\ln V_k - \ln \bar{V}|^{-\rho}\right]^{-\frac{1}{\rho}} . \qquad (5.8)$$

where:

$$\ln \bar{V} = \sum_{k=1}^{K} a_k(x) \ln V_k \left[\frac{p\, m(A_k)}{M_k}\right] .$$

We can complete the specification of a social welfare function W(u,x) by choosing a set of weights $a_k(x)$ for the levels of individual welfare $\left\{\ln V_k \left[\frac{p\, m(A_k)}{M_k}\right]\right\}$ in (5.8).

For this purpose we must appeal to a notion of vertical equity. Following Hammond [1977], we define a distribution of total expenditure $\{M_k\}$ as more **equitable** than another distribution $\{M_k'\}$ if:

(i) $M_i + M_j = M_i' + M_j'$,

(ii) $M_k = M_k'$ for $k \neq i, j$,

(iii) $\ln V_i \left[\frac{p\, m(A_i)}{M_i'}\right] > \ln V_i \left[\frac{p\, m(A_i)}{M_i}\right] > \ln V_j \left[\frac{p\, m(A_j)}{M_j}\right] > \ln V_j \left[\frac{p\, m(A_j)}{M_j'}\right]$,

We say that a social welfare function W(u,x) is **equity-regarding** if it is larger for a more equitable distribution of total expenditure.

We require that the social welfare function (5.8) must be equity-regarding. This amounts to imposing a version of Dalton's [1920] principle of transfers. This principle requires that a transfer of total expenditures from a rich household to a poor household that does not reverse their relative positions in the distribution of total expenditure must increase the level of social welfare.

If the social welfare function (5.8) is required to be equity-regarding, then the weights $\{a_k(x)\}$ associated with the individual welfare functions $\{\ln V_k[\frac{p\, m(A_k)}{M_k}]\}$ must take the form:

$$a_k(x) = \frac{m_0(p,A_k)}{\sum_{k=1}^{K} m_0(p,A_k)}, \quad (k = 1, 2 \ldots K). \tag{5.9}$$

We conclude that an equity-regarding social welfare function of the class (5.8) must take the form:

$$W(u,x) = \ln \bar{V} - \gamma(x)\left[\frac{\sum_{k=1}^{K} m_0(p,A_k) | \ln V_k - \ln \bar{V}|^{-\rho}}{\sum_{k=1}^{K} m_0(p,A_k)}\right]^{-\frac{1}{\rho}}, \tag{5.10}$$

where:

$$\ln \bar{V} = \frac{\sum_{k=1}^{K} m_0(p,A_k) \ln V_k\left[\frac{p\, m(A_k)}{M_k}\right]}{\sum_{k=1}^{K} m_0(p,A_k)},$$

$$= \ln p' \left(\alpha_p + \frac{1}{2} B_{pp} \ln p\right) - D(p) \frac{\sum_{k=1}^{K} m_0(p,A_k) \ln [M_k/m_0(p,A_k)]}{\sum_{k=1}^{K} m_0(p,A_k)}.$$

Furthermore, the condition of positive association implies that the function $\gamma(x)$ in (5.10) must take the form:

$$\gamma(x) = \left\{1 + \left[\frac{\sum_{k=1}^{K} m_0(p,A_k)}{m_0(p,A_j)}\right]^{-(\rho+1)}\right\}^{\frac{1}{\rho}}, \qquad (5.11)$$

where:

$$m_0(p,A_j) = \min_k m_0(p,A_k), \quad (k = 1, 2 \ldots K).$$

In order to formulate a social cost-of-living index, we can introduce the social expenditure function, defined as the minimum level of aggregate expenditure $M = \sum_{k=1}^{K} M_k$ required to attain a given level of social welfare, say W, at a specified price system p.[22] More formally, the social expenditure function M(p, W) is defined by:

$$M(p,W) = \min\left\{M{:}W(u,x) \geq W;\ M = \sum_{k=1}^{K} M_k\right\}. \qquad (5.12)$$

The social expenditure function (5.12) is precisely analogous to the individual expenditure function (2.19). The individual expenditure function gives the minimum level of individual expenditure required to attain a stipulated level of individual welfare; the social expenditure function gives the minimum level of aggregate expenditure required to attain a stipulated level of social welfare. Just as the individual expenditure function and the indirect utility function can be employed in measuring the individual cost of living, the social expenditure function and the social welfare function can be employed in measuring the social cost of living.

To construct a social expenditure function we first maximize social welfare for a fixed level of aggregate expenditure. We can maximize the average level of individual welfare

for a given level of aggregate expenditure by means of the Lagrangian:

$$Z = \ln \bar{V} + \lambda [\sum_{k=1}^{K} M_k - M], \tag{5.13}$$

$$= \ln p'(\alpha_p + \frac{1}{2} B_{pp} \ln p) - D(p) \frac{\sum_{k=1}^{K} m_0(p,A_k) \ln [M_k/m_0(p,A_k)]}{\sum_{k=1}^{K} m_0(p,A_k)}$$

$$+ \lambda [\sum_{k=1}^{K} M_k - M].$$

The first-order conditions for a constrained maximum of average individual welfare are:

$$\frac{D(p)}{\sum_{k=1}^{K} m_0(p,A_k)} \cdot \frac{m_0(p,A_k)}{M_k} = \lambda ,$$

$$\sum_{k=1}^{K} M_k = M, \qquad (k = 1, 2 \dots K),$$

so that total expenditure per household equivalent member $\{ M_k/m_0(p,A_k) \}$ is the same for all consuming units.

Next we consider the class of social welfare functions (5.8). Since the function g(x, u – $\bar{W}$ i) is nonpositive, we obtain a maximum of the social welfare function (5.8) if the function g(x, u – $\bar{W}$ i) can be made equal to zero while the average level of individual welfare $\bar{W}$ is a maximum. If total expenditure per household equivalent member $\{ M_k/m_0(p,A_k) \}$ is the same for all consuming units, the function g(x, u – $\bar{W}$ i) is equal to zero, so that the social welfare function W(u,x) in (5.8) is a maximum.

If aggregate expenditure is distributed so as to equalize total expenditure per household equivalent member, the level of individual welfare is the same for all consuming units. For this distribution of total expenditure the social welfare function (5.8) reduces to the average level of individual welfare $\ln \bar{V}$. For the translog indirect utility function the maximum value of social welfare for a given level of aggregate expenditure takes the form:

$$W(x,u) = \ln \bar{V}. \tag{5.14}$$

$$= \ln p'(\alpha_p + \frac{1}{2} B_{pp} \ln p) - D(p) \ln [M / \sum_{k=1}^{K} m_0(p,A_k)].$$

This value of social welfare is obtained by evaluating the translog indirect utility function (2.22) at total expenditure per household equivalent member $M / \sum_{k=1}^{K} m_0(p,A_k)$ for the economy as a whole. We can solve for aggregate expenditure as a function of the level of social welfare and prices:

$$\ln M(p, W) = \frac{1}{D(p)} [\ln p'(\alpha_p + \frac{1}{2} B_{pp} \ln p) - W] + \ln [\sum_{k=1}^{K} m_0(p,A_k)]. \tag{5.15}$$

We can refer to this function as the **translog social expenditure function**. The value of aggregate expenditure is obtained by evaluating the translog individual expenditure function (2.23) at the level of social welfare W and the number of household equivalent members $\sum_{k=1}^{K} m_0(p,A_k)$ for the economy as a whole.

To summarize: We have generated a class of social welfare functions (5.3) that has the properties of unrestricted domain, independence of irrelevant alternatives, positive association, nonimposition and cardinal full comparability. By imposing the additional assumption that the degree of aversion to inequality is constant and requiring the social welfare function to satisfy requirements of horizontal and vertical equity, we obtain the social welfare function (5.10). Finally, we have derived a social expenditure function (5.15) giving the minimum aggregate expenditure required to attain a base level of social welfare.

6. Social Cost-of-Living Index.

In this section we present methods for translating changes in prices into measures of change in the social cost of living. Following Pollak [1981], we define the social cost-of-living index, say P, as the ratio of two levels of aggregate expenditure:

$$P(p^1, p^0, W^0) = \frac{M(p^1, W^0)}{M(p^0, W^0)}. \tag{6.1}$$

In this ratio the numerator is the aggregate expenditure required to attain the base level of social welfare W^0 at the current price system p^1; the denominator is the expenditure required to attain this level of social welfare at the base price system p^0.

The social welfare function (5.10) and the translog social expenditure function (5.15) can be employed in implementing the social cost-of-living index (6.1). First, we can express the base level of social welfare W^0 in terms of the social welfare functions:

$$W^0 = \ln \bar{V}^0 - \tag{6.2}$$

$$\left\{1 + \left[\frac{\sum_{k=1}^{K} m_0(p^0,A_k)}{m_0(p^0,A_j)}\right]^{-1} (\rho - 1)\right\}^{\frac{1}{\rho}} \left[\frac{\sum_{k=1}^{K} m_0(p^0,A_k) | \ln V_k^0 - \ln \bar{V}^0|^{-\rho}}{\sum_{k=1}^{K} m_0(p^0,A_k)}\right]^{-\frac{1}{\rho}},$$

where:

$$\ln \bar{V}^0 = \frac{\sum_{k=1}^{K} m_0(p^0,A_k) \ln V_k \left[\frac{p^0\, m(A_k)}{M_k^0}\right]}{\sum_{k=1}^{K} m_0(p^0,A_k)},$$

and:

$$\ln V_k^0 = \ln p^{0\prime} (\alpha_p + \frac{1}{2} B_{pp} \ln p^0) - D(p^0) \ln [M_k^0/m_0(p^0, A_k)], (k = 1, 2 \dots K).$$

Using the translog social expenditure function (5.21), we can express the social cost-of-living index (6.1) in the form:

$$\ln P(p^1, p^0, W^0) = \frac{1}{D(p^1)} [\ln p^{1\prime} (\alpha_p + \frac{1}{2} B_{pp} \ln p^1) - W^0] \tag{6.3}$$

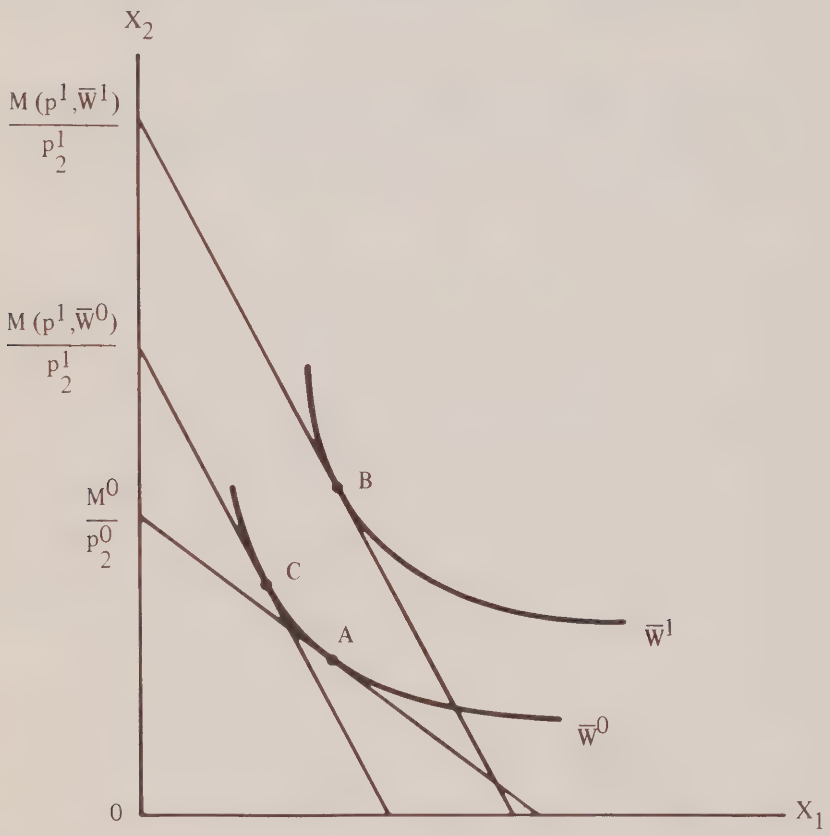
$$+ \ln [\sum_{k=1}^{K} m_0(p^1, A_k)] - \ln M^0.$$

We can refer to the index P as the **translog social cost-of-living index**. If the translog index is greater than unity and aggregate expenditure is constant, then social welfare is decreased by the change in prices.

The translog social expenditure function (5.15) has the same form as the translog individual expenditure function (2.23). We can express the level of social welfare as a function of the price system p and aggregate expenditure M. We can take this level of social welfare to be the average level of individual welfare $\bar{W}$, obtained by redistributing aggregate expenditure among households so as to equalize total expenditure per household equivalent member. Under this assumption society behaves in the same way as an individual maximizing a utility function, as demonstrated by Samuelson [1956] and Pollak [1981]. We can represent social welfare in terms of the indifference map for a single representative consumer, as in Diagram 1.

In Diagram 2 we have depicted the indifference map of the representative consumer with indirect utility function given by the average level of utility $\ln \bar{V}$ in (5.14). As before, we consider the case of two commodities (N = 2) for simplicity. Equilibrium of the representative consumer at the base price system p^0 is represented by the point A with base level of average welfare $\bar{W}^0$. The corresponding level of aggregate expenditure M^0, divided by the base price of the second commodity p_2^0, is given on the vertical axis. This axis provides a representation of aggregate expenditure in terms of units of the second commodity.

Diagram 2. Social Cost of Living Index.



Equilibrium of the representative consumer after the change in prices is represented by the point B with associated level of average welfare $\bar{W}^1$. The level of aggregate expenditure associated with the change in prices M^1, divided by the current price of the second commodity p_2^1, is given, as before, on the vertical axis. Finally, the level of aggregate expenditure required to attain the base level of average welfare $\bar{W}^0$ at current prices $M(p^1, \bar{W}^0)$, corresponds to equilibrium of the representative consumer at the point C. The social cost-of-living index is given by the ratio of the distances on the vertical axis corresponding to consumer equilibrium at points C and A, multiplied by the ratio of prices of the second commodity at the two points p_2^1 / p_2^0.

As a further illustration of the social cost-of-living index, we analyze the changes in prices over the period 1958-1978, using prices for 1972 as the base price system. As before, we employ the econometric model of aggregate consumer behavior presented in Section 3 for this purpose. Using the translog social cost-of-living index (6.3), we can assess the impact of price changes on the U.S. economy as a whole. We present the social cost-of-living index and rates of inflation corresponding to this index in Table 3. Over the 20-year period 1958-1978 this index has risen from .6928 to 1.5214 with 1972 equal to 1.0000.

To summarize: We have defined the social cost-of-living index as the ratio of the aggregate expenditure required to attain a base level of social welfare at current prices to the base level of expenditure. Using the average level of social welfare (5.14) and the translog social expenditure function (5.15), we implement this definition by means of the translog social cost-of-living index (6.3). We illustrate the translog index by analyzing changes in prices over the period 1958-1978 for the U.S. economy as a whole.

TABLE 3. Social Cost-of-Living Index

Year	Social Cost-of-Living Index (1972 = 1.0000)	Inflation Rate (annual percentage rates)
1958	0.6928	0.00
1959	0.6946	0.25
1960	0.7156	2.96
1961	0.7217	0.85
1962	0.7296	1.09
1963	0.7335	0.53
1964	0.7497	2.17
1965	0.7731	3.07
1966	0.7947	2.75
1967	0.8030	1.03
1968	0.8303	3.34
1969	0.8831	6.16
1970	0.8972	1.58
1971	0.9344	4.06
1972	1.0000	6.77
1973	1.0748	7.21
1974	1.1619	7.79
1975	1.2194	4.83
1976	1.3101	7.17
1977	1.4281	8.61
1978	1.5214	6.33

7. Group Cost-of-Living Indexes

In Sections 4 and 6 we have presented cost-of-living indexes for individual households and for the U.S. economy as a whole. In this section our objective is to provide measures of the cost of living for groups of households. For this purpose we introduce group welfare and expenditure functions that are precisely analogous to the social welfare and expenditure functions of Section 5. We consider a group of G individuals, where $1 \leq G \leq K$;

without loss of generality we can take the group to be comprised of the first G individuals in society.

To describe group orderings we find it useful to introduce the following notation:

x_{ng} - the quantity of the nth commodity group consumed by the gth consuming unit ($n = 1, 2 \ldots N$; $g = 1, 2 \ldots G$).

x_G - a matrix with elements $\{x_{ng}\}$ describing the group state.

$u_G = (W_1, W_2 \ldots W_G)$ - a vector of individual welfare functions for all G individuals.

We can define a group welfare functional as a mapping from the set of individual welfare functions to the set of group orderings. We can describe a group ordering in terms of properties of a group welfare functional that are precisely analogous to the properties of a social welfare functional considered in Section 5: unrestricted domain, independence of irrelevant alternatives, positive association, nonimposition and cardinal full comparability. Under these assumptions there exists a group ordering for the set of all group states that can be represented by a group welfare function analogous to the social welfare function (5.3):

$$W_G(u_G, x_G) = F\left\{\bar{W}_G(x_G) + g[x_G, u_G(x_G) - \bar{W}_G(x_G)\, i], x_G\right\}, \tag{7.1}$$

where the function $\bar{W}_G(x_G)$ corresponds to average individual welfare:

$$\bar{W}_G(x_G) = \sum_{g=1}^{G} a_g(x_G)\, W_g(x_G).$$

As before, the function g is a linear homogeneous function of deviations of levels of individual welfare from average welfare for the group.

We can rule out the direct dependence of the function $F(x_G)$ in (7.1) on characteristics of groups states x_G. As in Section 5, this results in a cardinal representation of group welfare. To complete the selection of a group welfare function $W_G(u_G, x_G)$ we require,

as before, that the individual welfare functions $\{W_k(x_G)\}$ must be invariant with respect to any positive affine transformation that is the same for all households. Second, we require that the group welfare function is equity-regarding. Under these assumptions the group welfare function must take forms analogous to (5.10).

In order to formulate a group cost-of-living index, we can introduce a group expenditure function, defined as the minimum level of group expenditure $M_G = \sum_{g=1}^{G} M_g$ required to attain a given level of group welfare, say W_G, at a specified price system p. We can maximize group welfare for a fixed level of group expenditure by equalizing total expenditure per household equivalent member. For the translog indirect utility function the group welfare function takes the form:

$$W_G(u_G, x_G) = \ln \bar{V}_G , \tag{7.2}$$

$$= \ln p' \left(\alpha_p + \frac{1}{2} B_{pp} \ln p\right) - D(p) \ln[M_G / \sum_{g=1}^{G} m_0(p, A_g)],$$

where $\ln \bar{V}_G$ is the average level of individual welfare.

We can solve for group expenditure as a function of the level of group welfare and prices:

$$\ln M_G(p, W_G) = \frac{1}{D(p)} [\ln p' (\alpha_p + \frac{1}{2} B_{pp} \ln p) - W_G] + \ln [\sum_{g=1}^{G} m_0(p, A_g)]. \tag{7.3}$$

We can refer to this function as the **translog group expenditure function**. The value of group expenditure is obtained by evaluating the translog individual expenditure function (2.23) at the level of group welfare W_G and the number of household equivalent members $\sum_{g=1}^{G} m_0(p, A_g)$ for the group as a whole.

We can define the group cost-of-living index, say P_G, as the ratio of two levels of group expenditure:

$$P_G(p^1, p^0, W_G^0) = \frac{M_G(p^1, W_G^0)}{M_G(p^0, W_G^0)}. \tag{7.4}$$

In this ratio the numerator is the group expenditure required to attain the base level of group welfare W_G^0 at the current price system p^1; the denominator is the expenditure required for this level of welfare at the base price system p^0. Using the translog group expenditure function (7.3), we can express the group cost-of-living index (7.4) in the form:

$$\ln P_G(p^1, p^0, W_G^0) = \frac{1}{D(p^1)}[\ln p^{1\prime}(\alpha_p + \frac{1}{2}B_{pp} \ln p^1) - W_G^0] \tag{7.5}$$

$$+ \ln[\sum_{g=1}^{G} m_0(p^1, A_g)] - \ln M_G^0 .$$

We can refer to the index P_G as the **translog group cost-of-living index**. If the translog index is greater than unity and group expenditure is constant, then group welfare is decreased by the change in prices.

The translog group expenditure function (7.3) has the same form as the translog social expenditure function (5.15). To obtain the group expenditure function from the social expenditure function, we replace social welfare W and the number of household equivalent members for society as a whole $\sum_{k=1}^{K} m_0(p, A_k)$ by group welfare W_G and the number of household equivalent members for the group $\sum_{g=1}^{G} m_0(p, A_g)$.

We can express the level of group welfare as a function of the price system p and group expenditure M_G. We can take this level of welfare to be the average level of individual welfare $\bar{W}_G$, obtained by equalizing total expenditure per household equivalent member

among all households in the group. Under this assumption the group behaves like an individual maximizing a utility function. We can represent group welfare in terms of the indifference map for a single representative consumer, just as we represented social welfare function in terms of such an indifference map in Diagram 2.

We can illustrate the group cost-of-living index by analyzing the changes in prices over the period 1958-1978, using prices for 1972 as a base price system. We have calculated cost-of-living indexes for 21 different groups of households, classified by demographic characteristics. These demographic characteristics include size of household, age of head of household, region of residence, race and urban versus rural residence. For this purpose we set the base level of group welfare at the level the group attained in 1972. We present translog group cost-of-living indexes for the period 1958-1978 in Appendix Table 2.

We present rates of inflation calculated from the translog group cost-of-living indexes for seven family size groups in Table 4. In virtually every year the rates of inflation increase or decrease monotonically with family size. The most dramatic changes occur between unattached individuals, families of size one, and families of size two or more. For example, in 1975 unattached individuals had an inflation rate of only 3.65 percent, while households of size two experienced an inflation rate of 4.78 percent and households with seven or more members had an inflation rate of 5.31 percent. At the opposite extreme, individuals experienced an increase in cost of living of 7.98 percent in 1972, while households of size two had a rate of 6.82 percent and households of size seven or more had only 6.33 percent.

We present the impact of changes in prices on measures of the group cost of living for age groups in Table 4. We observe that rates of inflation increase or decrease monotonically with age up to the age group 55-64. In fact, the pattern of inflation rates for different age groups is similar to that for different family sizes. The reason for this is that there is a sizable correlation between family size and age of the head of household. Younger and older age groups are associated with smaller family sizes, while families in the middle of the age range are associated with larger sizes. Finally, we present rates of inflation for groups living in different regions, the two racial groups, and urban versus rural residents in Table 4. Regional differences and differences between urban and rural residents are substantial, while racial differences are not.

To summarize: We have generated group welfare functions that are analogous to the social welfare functions (5.10). Second, we have derived a group expenditure function (7.3), giving the minimum group expenditure required to attain a base level of group welfare. Finally, we have defined group cost-of-living indexes in terms of the ratio of the group expenditure required to attain a base level of group welfare at current prices to the base level of expenditure. We illustrate group cost-of- living indexes by comparing rates of inflation among 21 demographic groups for the period 1958-1978.

TABLE 4. Change in Group Cost-of-Living Indexes (annual percentage rates)

	Family Size						
Year	1	2	3	4	5	6	7+
1959	−0.16	0.26	0.31	0.32	0.34	0.33	0.34
1960	3.43	3.00	2.94	2.90	2.82	2.82	2.73
1961	0.79	0.85	0.85	0.85	0.87	0.87	0.88
1962	1.09	1.09	1.09	1.09	1.08	1.08	1.08
1963	0.21	0.51	0.55	0.57	0.62	0.62	0.67
1964	2.54	2.18	2.14	2.12	2.08	2.07	2.03
1965	3.63	3.09	3.02	2.98	2.92	2.93	2.87
1966	2.63	2.72	2.73	2.76	2.82	2.84	2.93
1967	0.40	1.02	1.10	1.13	1.18	1.17	1.22
1968	3.06	3.32	3.36	3.39	3.42	3.40	3.45
1969	6.76	6.19	6.12	6.08	6.00	6.00	5.92
1970	0.38	1.50	1.64	1.73	1.93	1.93	2.17
1971	4.07	4.09	4.09	4.07	4.03	4.01	3.96
1972	7.98	6.82	6.67	6.60	6.46	6.47	6.33
1973	7.68	7.15	7.08	7.10	7.18	7.25	7.36
1974	6.23	7.68	7.84	7.95	8.28	8.38	8.69
1975	3.65	4.78	4.91	4.99	5.14	5.14	5.31
1976	7.71	7.30	7.24	7.13	6.92	6.86	6.57
1977	9.50	8.72	8.60	8.49	8.31	8.29	8.05
1978	6.58	6.27	6.24	6.26	6.34	6.38	6.50

TABLE 4. Change in Group Cost-of-Living Indexes (annual percentage rates) - Continued

	Age of Head					
Year	16-24	25-34	35-44	45-54	55-64	65 +
1959	−0.27	0.02	0.30	0.37	0.39	0.21
1960	3.62	3.20	2.92	2.83	2.83	3.00
1961	0.76	0.81	0.85	0.87	0.87	0.85
1962	1.07	1.06	1.08	1.10	1.10	1.09
1963	0.07	0.36	0.57	0.62	0.63	0.50
1964	2.63	2.36	2.16	2.07	2.05	2.19
1965	3.82	3.34	2.99	2.92	2.91	3.13
1966	2.50	2.66	2.74	2,82	2.81	2.77
1967	0.21	0.73	1.13	1.19	1.21	0.96
1968	2.93	3.21	3.42	3.40	3.40	3.28
1969	6.94	6.46	6.13	6.00	5.97	6.20
1970	−0.12	0.91	1.68	1.95	1.97	1.51
1971	4.10	4.03	4.06	4.06	4.07	4.06
1972	8.35	7.36	6.64	6.45	6.42	6.89
1973	7.71	7.41	7.05	7.15	7.12	7.36
1974	5.81	7.00	7.66	8.30	8.40	7.86
1975	3.25	4.21	4.92	5.16	5.20	4.74
1976	8.08	7.47	7.18	6.97	6.98	7.14
1977	9.90	9.01	8.49	8.37	8.38	8.69
1978	6.55	6.44	6.25	6.30	6.27	6.41

TABLE 4. Change in Group Cost-of-Living Indexes (annual percentage rates) - Concluded

Year	Region				Race		Residence	
	NE	NC	S	W	White	Non-White	Urban	Rural
1959	0.13	0.29	0.40	0.12	0.26	0.19	0.23	0.62
1960	3.06	2.95	2.81	3.14	2.97	2.92	2.99	2.63
1961	0.84	0.85	0.87	0.83	0.85	0.84	0.85	0.88
1962	1.08	1.09	1.08	1.10	1.09	1.06	1.09	1.08
1963	0.46	0.54	0.63	0.43	0.53	0.54	0.51	0.75
1964	2.26	2.14	2.04	2.32	2.17	2.17	2.19	1.85
1965	3.21	3.03	2.88	3.25	3.07	3.07	3.09	2.63
1966	2.76	2.74	2.79	2.70	2.75	2.81	2.75	2.78
1967	0.87	1.08	1.24	0.83	1.04	1.01	1.00	1.53
1968	3.27	3.35	3.42	3.28	3.34	3.35	3.33	3.50
1969	6.30	6.12	5.95	6.40	6.17	6.14	6.19	5.67
1970	1.32	1.62	1.98	1.17	1.57	1.62	1.52	2.42
1971	4.03	4.08	4.05	4.10	4.07	3.96	4.06	4.09
1972	7.06	6.69	6.36	7.20	6.77	6.75	6.83	5.82
1973	7.42	7.14	7.07	7.26	7.19	7.41	7.24	6.78
1974	7.53	7.90	8.40	6.93	7.77	8.05	7.70	9.23
1975	4.56	4.90	5.23	4.40	4.82	4.85	4.77	5.74
1976	7.19	7.21	6.98	7.44	7.20	6.92	7.19	6.93
1977	8.76	8.60	8.32	8.94	8.63	8.42	8.65	8.08
1978	6.46	6.27	6.25	6.35	6.31	6.50	6.35	6.04

8. Summary and Conclusion

In this paper we have presented an econometric approach to individual, group and social cost-of-living measurement. The key to our approach is provided by the translog indirect utility function (2.22) and the translog individual expenditure function (2.23). In Section 2 we show how the translog indirect utility and expenditure functions can be determined from a system of individual expenditure shares (2.5) that is integrable. We also demonstrate how the individual expenditure shares (2.5) can be recovered uniquely from the system of aggregate expenditure shares (2.25).

In Section 3 we fit an econometric model of aggregate consumer behavior that incorporates the restrictions implied by integrability of the individual expenditure shares. In

Section 4 we define a cost-of-living index for the individual consuming unit. We implement this index for the period 1958-1978, using translog indirect utility and expenditure functions for all consuming units.

We define and implement a cost-of-living index for the U.S. economy as a whole for the period 1958-1978 in Section 6. This definition is precisely analogous to the definition of an individual cost-of-living index. The role of the indirect utility function is played by an explicit social welfare function introduced in Section 5. The role of the individual expenditure function is played by the translog social expenditure function (5.15). This expenditure function is simply the translog individual expenditure function, evaluated at aggregate expenditure per household equivalent member.

In Section 7 we extend the concept of a social cost-of-living index to groups of individuals. We obtain a translog group cost-of-living index that can be expressed in terms of group welfare and expenditure functions. The group expenditure function is the translog individual expenditure function, evaluated at group expenditure per household equivalent member. We present cost-of-living indexes for 21 demographic groups for the period 1958-1978.

Next, we find it useful to compare the information required for implementation of the econometric approach to cost-of-living measurement with that required for conventional cost-of-living index numbers.[23] Our cross-section data set, the Survey of Consumer Expenditures for 1972, was assembled for the purpose of providing weights for the Consumer Price Index in the United States. We have employed this data set in estimating the impact of total expenditures and demographic characteristics of households on individual expenditure patterns.

Conventional cost-of-living index numbers are based on weighted averages of price relatives for individual commodity groups. The weights are based on averages of expenditure shares for groups of consumers. Measures of the cost-of-living for different groups can be constructed by compiling weights for each group. These weights require cross-section data like those collected in the Survey of Consumer Expenditures.

By contrast, the econometric approach is based on sample moments of cross-section data

on individual expenditures on specific commodities, total expenditures and the demographic characteristics of individual households. These sample moments are combined with time series data on prices and aggregate expenditure patterns to estimate the parameters of our econometric model. Estimates like those presented in Section 3 summarize the cross-section and time series data in a concise and readily intelligible way.

We have found it convenient to employ price data for commodity groups from the U.S. national accounts. These data are based largely on price information compiled for the Consumer Price Index published by the Bureau of Labor Statistics. Our approach could be implemented for price data taken directly from the Consumer Price Index. This would have the advantage of providing greater comparability between the results of the econometric approach and the index number approach employed currently.

Our econometric model employs time series data on the level of aggregate expenditure, its distribution over commodity groups, and its distribution over the population and among demographic groups. We have constructed these data from the Current Population Survey, which is conducted monthly in the United States. The necessary information is contained in the summary statistics of the joint distribution of expenditures and attributes -- $\Sigma M_k \ln M_k/M$ and $\{\Sigma M_k A_k/M\}$ -- presented in Section 2.

Although the econometric approach to cost-of-living measurement requires more data than the conventional index number approach, the econometric approach has overwhelming practical advantages. First and foremost, it allows for substitution among commodities in response to price changes. To achieve similar results by means of conventional index numbers, it would be necessary to have surveys of consumer expenditures at the same frequency as the intervals used for reporting the cost-of-living index, which would be far too burdensome.

A second advantage of the econometric approach to cost-of-living measurement is in flexibility and ease of application. We have implemented a social cost-of-living index for the United States in Section 6. In addition, we have compared cost-of-living increases for typical consuming units in Section 4. Finally, we have made comparisons among cost-of-living increases for groups of consuming units in Section 7. Our approach could also be

used to construct special cost-of-living indexes for groups such as low income households receiving government transfer payments or aged households living on pensions.

While the econometric approach has important advantages over the conventional index number approach, it is important to emphasize that the two approaches share a number of limitations. In principle prices for commodity groups must be compiled for goods and services of constant quality. In actuality statisticians are faced with rapidly changing commodity specifications and with the introduction of new commodities. As an illustration, the addition of pollution control equipment to automobiles poses exactly the same problems for the two approaches.

A second set of problems common to the econometric and the index number approaches is posed by consumption not associated with direct purchases of goods and services. As an illustration, the imputed value of home produced goods such as food produced on farms must be included in total expenditure. Similarly, an imputation is required for flows of services from owner-occupied housing and owner-utilized transportation equipment and other consumer durables. A more complex range of problems is posed by goods and services consumed collectively, such as police protection and environmental quality.

Finally, the implementation of the econometric approach described in this paper has important limitations of its own. We have employed prices for all commodity groups compiled at the national level. A more detailed implementation incorporating regional and other differences in prices actually paid would be useful in many applications. In addition, it would be very desirable to provide a more detailed commodity breakdown. These limitations can be overcome by expenditure of greater resources of human effort and computer time. Fortunately, existing data bases will be adequate for the construction of far more detailed econometric models than the model we have presented.

Up to this point we have compared the econometric approach with the conventional index number approach, such as that employed by the Bureau of Labor Statistics in the United States or by Statistics Canada. It is also interesting to compare the econometric approach with a more sophisticated index number approach proposed by Diewert [1976, 1981]. Diewert's approach is based on exact index numbers.[24] Exact index numbers do

not require an econometric model, but reproduce exactly the individual cost-of-living index derived from an individual expenditure function.

An important example of the sophisticated index number approach is the exact translog cost-of-living index considered by Tornqvist [1936]:

$$\ln P_k(p^1, p^0, V_k^*) = \bar{w}_k' \, \Delta \ln p, \qquad (k = 1, 2 \dots K), \tag{8.1}$$

where V_k^* is the base level of individual welfare:

$$V_k^* = (V_k^1 V_k^0)^{\frac{1}{2}}, \qquad (k = 1, 2 \dots K),$$

and:

$$\bar{w}_k = \frac{1}{2}(w_k^1 + w_k^0), \qquad (k = 1, 2 \dots K),$$

$$\Delta \ln p = \ln p^1 - \ln p^0.$$

Diewert [1976, 1981] has shown that this cost-of-living index is exact for a translog individual expenditure function similar but not identical to our expenditure function (2.23).

The data required for implementation of the exact index number approach are far more extensive than those required for the econometric approach. For example, the exact translog cost-of-living index (8.1) would require data on individual expenditure shares for each period in which a cost-of-living comparison is needed. An annual time series of comparisons would require that a panel of consumers would have to be surveyed annually.

It is important to note that the exact index number approach is not limited to individual cost-of-living measurement. For social or group cost-of-living measurement a translog cost-of-living index could be defined on the basis of the corresponding translog social or group expenditure functions. The individual budget shares in the index (8.1) could be replaced by aggregate or group expenditure shares.

In order to apply the exact translog cost-of-living index (8.1) to society as a whole, it would be necessary to obtain aggregate expenditure shares from an econometric model. These shares would correspond to those of a representative consumer, constructed by equalizing aggregate expenditure per household equivalent member over all consuming units. Data on aggregate expenditure shares employed in fitting an econometric model would not be appropriate for this purpose. These expenditure shares correspond to the actual distribution of total expenditure over the population.

Similarly, the exact translog cost-of-living index (8.1) could be applied to groups, but only by calculating expenditure shares for the group for an econometric model. As before, these shares would correspond to those of a representative consumer. The demand system for the representative consumer would be obtained by equalizing group expenditure per household equivalent member over the households included in the group.

Despite the limitations of the exact index number approach, we find it useful to present an exact translog social cost-of-living index calculated from the Tornqvist formula (8.1) in Table 5. This index can be compared with the translog social cost-of-living index presented in Table 3. The two indexes are similar, but not identical. An important difference between the two is that the exact translog index is a chain index with base welfare levels changing from period to period, while the translog index presented in Table 3 employs the level of welfare in 1972 as a base.

Our final conclusion is that the econometric approach has very substantial advantages over both conventional and sophisticated index number approaches. However, the econometric approach is not a panacea for the solution of all the practical problems of cost-of-living measurement. As better solutions to these problems become available, the results can be incorporated into an econometric model like that we have presented.

TABLE 5. Exact Translog Cost-of-Living Index

Year	Exact Cost-of-Living Index (1972 = 1.0000)	Inflation Rate ((annual percentage rates)
1958	.6934	----
1959	.6952	0.26
1960	.7156	2.89
1961	.7216	0.83
1962	.7293	1.06
1963	.7334	0.56
1964	.7490	2.10
1965	.7730	3.15
1966	.7939	2.67
1967	.8018	0.99
1968	.8295	3.40
1969	.8823	6.17
1970	.8981	1.77
1971	.9353	4.06
1972	1.0000	6.69
1973	1.0755	7.28
1974	1.1632	7.84
1975	1.2310	5.66
1976	1.3142	6.54
1977	1.4330	8.65
1978	1.5254	6.25

APPENDIX TABLE 1. Individual Cost-of-Living Indexes (1972 = 1.0000)

Year	Urban		Rural	
	White	Nonwhite	White	Nonwhite
1958	0.6969	0.6968	0.7002	0.7001
1959	0.6967	0.6964	0.7019	0.7016
1960	0.7176	0.7173	0.7209	0.7207
1961	0.7236	0.7233	0.7272	0.7269
1962	0.7311	0.7308	0.7346	0.7342
1963	0.7345	0.7341	0.7392	0.7389
1964	0.7508	0.7507	0.7535	0.7534
1965	0.7758	0.7756	0.7760	0.7758
1966	0.7992	0.7989	0.7998	0.7994
1967	0.8051	0.8048	0.8086	0.8084
1968	0.8310	0.8311	0.8353	0.8354
1969	0.8840	0.8845	0.8849	0.8854
1970	0.8973	0.8973	0.9042	0.9042
1971	0.9319	0.9317	0.9390	0.9388
1972	1.0000	1.0000	1.0000	1.0000
1973	1.0856	1.0848	1.0834	1.0826
1974	1.1864	1.1818	1.2014	1.1968
1975	1.2430	1.2375	1.2681	1.2625
1976	1.3251	1.3197	1.3479	1.3425
1977	1.4407	1.4341	1.4590	1.4524
1978	1.5459	1.5385	1.5633	1.5558

Region = Northeast.
Size = 5.
Age = 35-44.
Expenditure = $4,467.

APPENDIX TABLE 1. Individual Cost-of-Living Indexes (1972 = 1.0000) - Continued

Year	Urban		Rural	
	White	Nonwhite	White	Nonwhite
1958	0.6946	0.6945	0.6978	0.6978
1959	0.6952	0.6949	0.7004	0.7001
1960	0.7162	0.7159	0.7195	0.7193
1961	0.7222	0.7219	0.7258	0.7255
1962	0.7299	0.7296	0.7334	0.7330
1963	0.7335	0.7332	0.7382	0.7379
1964	0.7500	0.7498	0.7527	0.7525
1965	0.7742	0.7740	0.7745	0.7743
1966	0.7967	0.7964	0.7973	0.7970
1967	0.8037	0.8034	0.8072	0.8069
1968	0.8304	0.8305	0.8348	0.8349
1969	0.8836	0.8841	0.8845	0.8850
1970	0.8969	0.8970	0.9039	0.9039
1971	0.9328	0.9326	0.9398	0.9397
1972	1.0000	1.0000	1.0000	1.0000
1973	1.0802	1.0794	1.0780	1.0771
1974	1.1720	1.1675	1.1868	1.1823
1975	1.2282	1.2228	1.2530	1.2475
1976	1.3145	1.3092	1.3372	1.3318
1977	1.4309	1.4244	1.4491	1.4425
1978	1.5301	1.5228	1.5473	1.5399

Region = Northeast.
Size = 5.
Age = 35-44.
Expenditure = $8,934.

APPENDIX TABLE 1. Individual Cost-of-Living Indexes (1972 = 1.0000) - Concluded

Year	Urban		Rural	
	White	Nonwhite	White	Nonwhite
1958	0.6923	0.6922	0.6955	0.6954
1959	0.6937	0.6934	0.6988	0.6986
1960	0.7148	0.7145	0.7181	0.7179
1961	0.7209	0.7206	0.7244	0.7241
1962	0.7287	0.7284	0.7322	0.7318
1963	0.7326	0.7322	0.7373	0.7369
1964	0.7491	0.7490	0.7518	0.7517
1965	0.7727	0.7724	0.7729	0.7727
1966	0.7943	0.7939	0.7948	0.7945
1967	0.8023	0.8020	0.8058	0.8055
1968	0.8298	0.8299	0.8342	0.8343
1969	0.8833	0.8837	0.8842	0.8846
1970	0.8966	0.8966	0.9035	0.9035
1971	0.9336	0.9335	0.9407	0.9405
1972	1.0000	1.0000	1.0000	1.0000
1973	1.0748	1.0740	1.0725	1.0717
1974	1.1577	1.1533	1.1724	1.1679
1975	1.2137	1.2083	1.2382	1.2327
1976	1.3041	1.2988	1.3266	1.3212
1977	1.4212	1.4147	1.4392	1.4327
1978	1.5145	1.5073	1.5315	1.5242

Region = Northeast.
Size = 5.
Age = 35-44.
Expenditure = $17,864.

APPENDIX TABLE 2. Group Cost-of-Living Indexes (1972 = 1.0000)

	Family Size						
Year	1	2	3	4	5	6	7+
1958	0.6917	0.6927	0.6930	0.6930	0.6932	0.6935	0.6932
1959	0.6906	0.6946	0.6952	0.6953	0.6956	0.6958	0.6956
1960	0.7147	0.7157	0.7160	0.7158	0.7156	0.7157	0.7148
1961	0.7204	0.7219	0.7221	0.7219	0.7218	0.7220	0.7212
1962	0.7283	0.7298	0.7301	0.7299	0.7297	0.7299	0.7291
1963	0.7299	0.7336	0.7341	0.7341	0.7343	0.7344	0.7340
1964	0.7487	0.7498	0.7500	0.7499	0.7497	0.7498	0.7492
1965	0.7764	0.7733	0.7730	0.7726	0.7720	0.7722	0.7710
1966	0.7971	0.7947	0.7944	0.7942	0.7941	0.7945	0.7940
1967	0.8004	0.8029	0.8033	0.8033	0.8036	0.8038	0.8037
1968	0.8252	0.8300	0.8308	0.8311	0.8316	0.8317	0.8320
1969	0.8830	0.8831	0.8832	0.8832	0.8831	0.8831	0.8828
1970	0.8864	0.8965	0.8979	0.8987	0.9003	0.9004	0.9021
1971	0.9232	0.9340	0.9354	0.9361	0.9374	0.9373	0.9386
1972	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
1973	1.0798	1.0741	1.0734	1.0736	1.0744	1.0752	1.0764
1974	1.1493	1.1600	1.1610	1.1624	1.1672	1.1692	1.1742
1975	1.1921	1.2168	1.2195	1.2219	1.2288	1.2309	1.2382
1976	1.2877	1.3090	1.3111	1.3123	1.3169	1.3184	1.3224
1977	1.4161	1.4283	1.4290	1.4287	1.4310	1.4325	1.4333
1978	1.5125	1.5208	1.5210	1.5211	1.5247	1.5270	1.5295

APPENDIX TABLE 2. Group Cost-of-Living Indexes (1972 = 1.0000) - Continued

	Age of Head					
Year	16-24	25-34	35-44	45-54	55-64	65 +
1958	0.6930	0.6938	0.6926	0.6926	0.6928	0.6927
1959	0.6911	0.6940	0.6947	0.6952	0.6956	0.6943
1960	0.7166	0.7165	0.7153	0.7152	0.7156	0.7154
1961	0.7221	0.7224	0.7214	0.7215	0.7219	0.7216
1962	0.7299	0.7301	0.7293	0.7295	0.7299	0.7295
1963	0.7305	0.7328	0.7335	0.7341	0.7345	0.7332
1964	0.7500	0.7503	0.7495	0.7494	0.7497	0.7495
1965	0.7792	0.7758	0.7723	0.7717	0.7719	0.7734
1966	0.7990	0.7967	0.7938	0.7938	0.7939	0.7952
1967	0.8008	0.8026	0.8029	0.8034	0.8036	0.8028
1968	0.8246	0.8288	0.8308	0.8312	0.8314	0.8297
1969	0.8839	0.8842	0.8834	0.8826	0.8826	0.8828
1970	0.8828	0.8923	0.8984	0.9000	0.9003	0.8962
1971	0.9198	0.9290	0.9357	0.9374	0.9377	0.9333
1972	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
1973	1.0802	1.0769	1.0730	1.0741	1.0738	1.0763
1974	1.1448	1.1550	1.1585	1.1671	1.1680	1.1644
1975	1.1827	1.2048	1.2170	1.2290	1.2304	1.2210
1976	1.2823	1.2982	1.3076	1.3177	1.3194	1.3114
1977	1.4159	1.4207	1.4236	1.4328	1.4349	1.4306
1978	1.5119	1.5154	1.5154	1.5261	1.5278	1.5254

APPENDIX TABLE 2. Group Cost-of-Living Indexes (1972 = 1.0000) - Concluded

Year	NE	NC	S	W	White	Non-White	Urban	Rural
1958	0.6927	0.6932	0.6936	0.6910	0.6927	0.6942	0.6927	0.6953
1959	0.6937	0.6953	0.6964	0.6919	0.6945	0.6956	0.6943	0.6997
1960	0.7152	0.7161	0.7162	0.7140	0.7155	0.7162	0.7154	0.7184
1961	0.7213	0.7222	0.7225	0.7200	0.7216	0.7223	0.7215	0.7248
1962	0.7292	0.7302	0.7304	0.7280	0.7296	0.7300	0.7294	0.7327
1963	0.7325	0.7342	0.7351	0.7312	0.7335	0.7340	0.7332	0.7383
1964	0.7493	0.7501	0.7503	0.7484	0.7496	0.7501	0.7495	0.7521
1965	0.7738	0.7732	0.7722	0.7732	0.7730	0.7735	0.7731	0.7722
1966	0.7954	0.7947	0.7942	0.7944	0.7946	0.7956	0.7947	0.7940
1967	0.8024	0.8034	0.8041	0.8011	0.8029	0.8037	0.8028	0.8063
1968	0.8291	0.8308	0.8322	0.8278	0.8302	0.8311	0.8300	0.8351
1969	0.8831	0.8833	0.8833	0.8826	0.8830	0.8838	0.8831	0.8838
1970	0.8949	0.8978	0.9010	0.8930	0.8971	0.8983	0.8967	0.9055
1971	0.9317	0.9352	0.9383	0.9304	0.9344	0.9346	0.9339	0.9434
1972	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
1973	1.0770	1.0740	1.0733	1.0753	1.0745	1.0769	1.0750	1.0701
1974	1.1612	1.1623	1.1674	1.1525	1.1614	1.1672	1.1612	1.1737
1975	1.2154	1.2207	1.2302	1.2044	1.2188	1.2252	1.2179	1.2432
1976	1.3060	1.3119	1.3192	1.2974	1.3099	1.3130	1.3087	1.3324
1977	1.4256	1.4299	1.4336	1.4189	1.4281	1.4284	1.4270	1.4446
1978	1.5209	1.5225	1.5262	1.5119	1.5211	1.5244	1.5206	1.5346

Footnotes

[1] The literature on the individual cost-of-living index is summarized by Deaton and Muellbauer [1980], pp.170-178, and by Diewert [1981].

[2] The translog indirect utility function was introduced by Christensen, Jorgenson and Lau [1975] and extended to encompass determinants of expenditure allocation other than prices and total expenditure by Jorgenson and Lau [1975]. Alternative approaches to the representation of the effects of prices and total expenditure on expenditure allocation are summarized by Barten [1977], Deaton and Muellbauer [1980], pp.60-85, and Lau [1977a].

[3] Alternative approaches to the representation of household characteristics on expenditure allocation are presented by Barten [1964], Gorman [1976], and Prais and Houthakker [1971]. Empirical evidence on the impact of demographic characteristics on expenditure allocation is given by Lau, Lin and Yotopoulos [1978], Muellbauer [1977], Parks and Barten [1973], Pollak and Wales [1980, 1981], and Ray [1982].

[4] The specification of a system of individual demand functions by means of Roy's Identity was first employed in econometric modeling of consumer behavior by Houthakker [1960]. A detailed review of econometric models based on Roy's Identity is given by Lau [1977a].

[5] For further discussion, see Lau [1977b, 1982] and Jorgenson, Lau and Stoker [1980, 1981, 1982].

[6] Conditions for integrability are discussed by Jorgenson and Lau [1979] and by Jorgenson, Lau and Stoker [1982].

[7] For further discussion see Jorgenson, Lau and Stoker [1982], esp. pp.175-186.

[8] Household equivalence scales are discussed by Barten [1964], Lazear and Michael [1980], Muellbauer [1974, 1977, 1980], and Prais and Houthakker [1971], among others. Alternative approaches are summarized by Deaton and Muellbauer [1980].

[9] The use of household equivalence scales in evaluating transfers among individuals has been advocated by Deaton and Muellbauer [1980], esp. pp.205-212, and by Muellbauer [1974]. Pollak and Wales [1979] have presented arguments against the use of household equivalence scales for this purpose.

[10] The 1972-1973 Survey of Consumer Expenditures is discussed by Carlson [1974].

[11] We employ data on the flow of services from durable goods rather than purchases of durable goods. Personal consumption expenditures in the U.S. National Income and Product Accounts are based on purchases of durable goods.

[12] This series is published annually by the U.S. Bureau of the Census.

[13] A detailed discussion of the stochastic specification of our model and of econometric methods for pooling time series and cross-section data is presented by Jorgenson, Lau and Stoker [1982], Section 6. This stochastic specification implies that time series data must be adjusted for

heteroscedasticity by multiplying each observation by the statistic:

$$\rho = \frac{(\Sigma\, M_k)^2}{\Sigma\, M_k^2}.$$

14 An alterative approach to implementation of the individual cost-of- living index (4.2) is to bound this index on the basis of observable data. Bounds have been developed by Allen [1949], Frisch [1936] and Konüs [1939] and, more recently, by Afriat [1977], Pollak [1971, 1981], and Samuelson and Swamy [1974], among many others. This approach is reviewed by Deaton and Muellbauer [1980], pp.170-178, and by Diewert [1981].

15 Arrow [1977, p.225] has defended noncomparability in the following terms: ... the autonomy of individuals, an element of mutual incommensurability among people seems denied by the possibility of interpersonal comparisons.

16 It is important to note that the social welfare function is (5.2) represents a social ordering over all possible individual orderings and exemplifies the multiple profile approach to social choice Arrow [1963] rather than the single profile approach employed by Bergson [1938] and Samuelson [1947]. The literature on the existence of single profile social welfare functions is discussed by Roberts [1980d], Samuelson [1982] and Sen [1979b].

17 See Sen [1977, 1979b] for further discussion.

18 See Roberts [1980b], esp. pp.434-436.

19 The implications of distributional homotheticity are discussed by Kolm [1976b] and Blackorby and Donaldson [1978].

20 Mean value functions were introduced into economics by Bergson [1936] and have been employed, for example, by Arrow, Chenery, Minhas and Solow [1961] and Atkinson [1970]. Properties of mean value functions are discussed by Hardy, Littlewood and Polya [1959].

21 This assumption implies that individual welfare increases with total expenditure at a rate that is inversely proportional to total expenditure. This is also implied by the utilitarian social welfare function employed by Arrow and Kalt [1979].

22 The social expenditure function was introduced by Pollak [1981]. Alternative money measures of social welfare are discussed by Arrow and Kalt [1979], Bergson [1980], Deaton and Muellbauer [1980], pp.214-239, Roberts [1980c], and Sen [1976]. A survey of the literature is presented by Sen [1979a].

23 The conventional index number approach is summarized in Statistics Canada [1982].

24 Exact index numbers are discussed by Lau [1979], Pollak [1971] and Samuelson and Swamy [1974].

References

Afriat, S.N. [1977], *The Price Index*, London, Cambridge University Press.

Allen, R.G.D. [1949], ''The Economic Theory of Index Numbers'', *Economica*, Vol. 16, No. 63, August, pp.197-203.

Arrow, K.J. [1963], *Social Choice and Individual Values*, New Haven, Yale University Press, 2nd ed.

__________ [1977], ''Extended Sympathy and the Possibility of Social Choice'', *American Economic Review*, Vol. 67, No. 1, February, pp.219-225.

__________ H.B. Chenery, B.S. Minhas and R.M. Solow [1961], ''Capital-Labor Substitution and Economic Efficiency'', *Review of Economics and Statistics*, Vol. 43, No. 3, August, pp.225-250.

__________ and J.P. Kalt [1979], *Petroleum Price Regulation*: *Should We Decontrol*, Washington, American Enterprise Institute.

d'Aspremont, C. and L. Gevers [1977], ''Equity and the Informational Basis of Collective Choice'', *Review of Economic Studies*, Vol. 44, No. 137, June, pp.199-209.

Atkinson, A.B. [1970], ''On Measurement of Inequality'', *Journal of Economic Theory*, Vol. 2, No. 3, September, pp.244-263.

Barten, A.P. [1964], ''Family Composition, Prices, and Expenditure Patterns'', in P. Hart, G. Mills and J.K. Whitaker (eds.), *Econometric Analysis for National Economic Planning*: *16th Symposium of the Colston Society*, London, Butterworth, pp.277-292.

__________ [1977], ''The Systems of Consumer Demand Functions Approach: A Review'', in M.D. Intriligator (ed.), *Frontiers of Quantitative Economics*, Vol. IIIA, Amsterdam, North-Holland, pp.23-58.

Bergson, A. [1936], "Real Income, Expenditure Proportionality, and Frisch's 'New Methods of Measuring Marginal Utility'", *Review of Economic Studies*, Vol. 4, No. 1, October 1936, pp.33-52.

_________ [1938], "A Reformulation of Certain Aspects of Welfare Economics", *Quarterly Journal of Economics*, Vol. 52, No. 2, February, pp.310-334.

_________ [1980], "Consumer's Surplus and Income Redistribution", *Journal of Public Economics*, Vol. 14, No. 1, August, pp.31-47.

Blackorby, C. and D. Donaldson [1978], "Measures of Relative Equality and their Meaning in Terms of Social Welfare", *Journal of Economic Theory*, Vol. 18, No. 1, June, pp.651-675.

_________ [1982], "Ratio-Scale and Translation-Scale Full Interpersonal Comparability without Domain Restrictions: Admissible Social-Evaluation Functions", *International Economic Review*, Vol. 23, No. 2, June, pp.249-268.

Bureau of the Census (various annual issues), *Current Population Reports, Consumer Income, Series P-60*, Washington, D.C., U.S. Department of Commerce.

Carlson, M.D. [1974], "The 1972-73 Consumer Expenditure Survey", *Monthly Labor Review*, Vol. 97, No. 12, December, pp.16-23.

Christensen, L.R., D.W. Jorgenson and L.J. Lau [1975], "Transcendental Logarithmic Utility Functions", *American Economic Review*, Vol. 65, No. 3, June, pp.367-383.

Dalton, H. [1920], "The Measurement of Inequality of Income", *Economic Journal*, Vol. 30, No. 119, September, pp.361-384.

Deaton, A. and J. Muellbauer [1980], *Economics and Consumer Behavior*, Cambridge, Cambridge University Press.

Deschamps, R. and L. Gevers [1978], ''Leximin and Utilitarian Rules: A Joint Characterisation'', *Journal of Economic Theory*, Vol. 17, No. 2, April, pp.143-163.

Diewert, W.E. [1976], ''Exact and Superlative Index Numbers'', *Journal of Econometrics*, Vol. 4, No. 2, May, pp.115-145.

__________ [1981], ''The Economic Theory of Index Numbers: A Survey'', in A. Deaton (ed.), *Essays on Theory and Measurement of Consumer Behavior in Honor of Sir Richard Stone*, pp.163-208.

Frisch, R. [1936], ''Annual Survey of Economic Theory: The Problem of Index Numbers'', *Econometrica*, Vol. 4, No. 1, January, pp.1-39.

Gorman, W.M. [1976], ''Tricks with Utility Functions'', in M.J. Artis and A.R. Nobay (eds.), *Essays in Economic Analysis*: *Proceedings of the 1975 AUTE Conference*, Cambridge, Cambridge University Press, pp.211-243.

Hammond, P.J. [1976], ''Equity, Arrow's Conditions and Rawl's Difference Principle'', *Econometrica*, Vol. 44, No. 4, July, pp.793-804.

__________ [1977], ''Dual Interpersonal Comparisons of Utility and the Welfare Economics of Income Distribution'', *Journal of Public Economics*, Vol. 7, No. 1, February, pp.51-71.

Hardy, G.H., J.E. Littlewood and G. Polya [1959], *Inequalities*, Cambridge, Cambridge University Press, 2nd ed.

Harsanyi, J.C. [1976], *Essays on Ethics, Social Behavior and Scientific Explanation*, Dordrecht, D. Reidel.

Houthakker, H.S. [1960], ''Additive Preferences'', *Econometrica*, Vol. 28, No. 2, April, pp.244-257.

Jorgenson, D.W. and L.J. Lau [1975], "The Structure of Consumer Preferences", *Annals of Economic and Social Measurement*, Vol. 4, No. 1, January, pp.49-101.

_________ and _________ [1979], "The Integrability of Consumer Demand Functions", *European Economic Review*, Vol. 12, No. 2, April, pp.115-147.

_________ L.J. Lau and T.M. Stoker [1980], "Welfare Comparison Under Exact Aggregation", *American Economic Review*, Vol. 70, No. 2, May, pp.268-272.

_________, _________, and _________ [1981], "Aggregate Consumer Behavior and Individual Welfare", in D. Currie, R. Nobay and D. Peel (eds.), *Macroeconomic Analysis*, London, Croom-Helm, pp.35-61.

_________ [1982], "The Transcendental Logarithmic Model of Aggregate Consumer Behavior", in R.L. Basmann and G.F. Rhodes Jr. (eds.), *Advances in Econometrics*, Vol. 1, Greenwich, JAI Press, pp.97-238.

Kolm, S.C. [1969], "The Optimal Production of Social Justice", in J. Margolis and H. Guitton (eds.), *Public Economics*, London: Macmillan, pp.145-200.

_________ [1976a and 1976b], "Unequal Inequalities I and II", *Journal of Economic Theory*, Vol. 12, No. 3, June, pp.416-42, and Vol. 13, No. 1, August, pp.82-111.

Konüs, A.A. [1939], "The Problem of the True Index of the Cost-of-Living", *Econometrica*, Vol. 7, No. 1, January, pp.10-29; originally published in 1924.

Lau, L.J. [1977a], "Complete Systems of Consumer Demand Functions through Duality", in M.D. Intriligator (ed.), *Frontiers of Quantitative Economics*, Vol. IIIA, Amsterdam, North-Holland, pp.59-86.

_________ [1977b], "Existence Conditions for Aggregate Demand Functions", Technical Report No. 248, Institute for Mathematical Studies in the Social Sciences, Stanford University, Stanford (revised 1980 and 1982).

_________ [1979], "On Exact Index Numbers", *Review of Economics and Statistics*, Vol. 61, No. 1, February, pp.73-82.

_________ [1982], "A Note on the Fundamental Theorem of Exact Aggregation", *Economics Letters*, Vol. 9, No. 2, pp.119-126.

Lau, L.J., W.L. Lin and P.A. Yotopoulos [1978], "The Linear Logarithmic Expenditure System: An Application to Consumption-Leisure Choice", *Econometrica*, Vol. 46, No. 4, July, pp.843-868.

Lazear, E.P. and R.T. Michael [1980], "Family Size and the Distribution of Real Per Capita Income", *American Economic Review*, Vol. 70, No. 1, March, pp.91-107.

Martos, B. [1969], "Subdefinite Matrices and Quadratic Forms", *SIAM Journal of Applied Mathematics*, Vol. 17, pp.1215-1223.

Maskin, E. [1978], "A Theorem on Utilitarianism", *Review of Economic Studies*, Vol. 42, No. 139, February, pp.93-96.

Mirrlees, J.A. [1971], "An Exploration in the Theory of Optimal Income Taxation", *Review of Economic Studies*, Vol. 38, No. 114, April, pp.175-208.

Muellbauer, J. [1974], "Household Composition, Engel Curves and Welfare Comparisons between Households: A Duality Approach", *European Economic Review*, Vol. 5, No. 2, August, pp.103-122.

_________ [1977], "Testing the Barten Model of Household Composition Effects and the Cost of Children", *Economic Journal*, Vol. 87, No. 347, September, pp. 460-487.

_________ [1980], "The Estimation of the Prais-Houthakker Model of Equivalence Scales", *Econometrica*, Vol. 48, No. 1, January, pp. 153-176.

Ng, Y.K. [1975], "Bentham or Bergson? Finite Sensibility, Utility Functions and Social Welfare Functions", *Review of Economic Studies*, Vol 42, No. 4, October, pp. 545-569.

Parks, R.W. and A.P. Barten [1973], "A Cross Country Comparison of the Effects of Prices, Income, and Population Composition on Consumption Patterns", *Economic Journal*, Vol. 83, No. 331, September, pp. 834-852.

Pollak, R.A. [1971], "The Theory of the Cost of Living Index", Research Discussion Paper No. 11, Office of Prices and Living Conditions, U.S. Bureau of Labor Statistics, Washington, D.C.

__________ [1981], "The Social Cost of Living Index", *Journal of Public Economics*, Vol. 15, No. 3, June, pp. 311-336.

Pollak, R.A. and T.J. Wales [1979], "Welfare Comparisons and Equivalent Scales", *American Economic Review*, Vol. 69, No. 2, May, pp. 216-21.

__________ , __________ , and __________ [1980] "Comparisons of the Quadratic Expenditure System and Translog Demand Systems with Alternative Specifications of Demographic Effects", *Econometrica*, Vol. 48, No. 3, April, pp. 595-612.

__________ , __________ , and __________ [1981] "Demographic Variables in Demand Analysis", *Econometrica*, Vol. 49, No. 6, November, pp. 1533-1552.

Prais, S.J. and H.S. Houthakker [1971], *The Analysis of Family Budgets*, Cambridge, Cambridge University Press, 2nd ed.

Ray, R. [1982], "The Testing and Estimation of Complete Demand Systems on Household Budget Surveys: An Application of AIDS", *European Economic Review*, Vol. 17, No. 3, March 1982, pp. 349-370.

Roberts, K.W.S. [1980a], "Possibility Theorems with Interpersonally Comparable Welfare Levels", *Review of Economic Studies*, Vol. 47, No. 147, January, pp. 409-20.

__________ [1980b], "Interpersonal Comparability and Social Choice Theory", *Review of Economic Studies*, Vol. 47, No. 147, January, pp. 421-439.

________ [1980c], ''Price-Independent Welfare Prescriptions'', *Journal of Public Economics*, Vol. 13, No. 3, June, pp. 277-298.

________ [1980d], ''Social Choice Theory: The Single-profile and Multi-profile Approaches'', *Review of Economic Studies*, Vol. 47, No. 147, January, pp. 441-450.

Roy, R. [1943], *De l'Utilite: Contributions a la Theorie des Choix*, Paris, Herman.

Samuelson, P.A. [1947], *Foundations of Economic Analysis*, Cambridge, Harvard University Press.

________ [1956], ''Social Indifference Curves'', *Quarterly Journal of Economics*, Vol. 70, No. 1, February, pp. 1-22.

________ [1982], ''Bergsonian Welfare Economics'', in S. Rosefielde (ed.), *Economic Welfare and the Economics of Soviet Socialism: Essays in Honor of Abram Bergson*, Cambridge, Cambridge University Press, pp. 223-266.

________ and S. Swamy [1974], ''Invariant Economic Index Numbers and Canonical Duality: Survey and Synthesis'', *American Economic Review*, Vol. 64, No. 4, September, pp. 566-93.

Sen, A.K. [1970], *Collective Choice and Social Welfare*, Edinburgh, Oliver and Boyd.

________ [1973], *On Economic Inequality*, Oxford, Clarendon Press.

________ [1976], ''Real National Income'', *Review of Economic Studies*, Vol. 43, No. 133, February, pp.19-40.

________ [1977], ''On Weights and Measures: Informational Constraints in Social Welfare Analysis'', *Econometrica,* Vol. 45, No. 7, October, pp.1539-1572.

_________ [1979a], ''The Welfare Basis of Real Income Comparisons: A Survey'', *Journal of Economic Literature*, Vol. 17, No. 1, March, pp.1-45.

_________ [1979b], ''Personal Utilities and Public Judgements: Or What's Wrong with Welfare Economics'', *Economic Journal*, Vol. 89, No. 763, September, pp.537-558.

Statistics Canada [1982], *The Consumer Price Index Reference Paper*, Ottawa, Statistics Canada, May.

Tornqvist, L. [1936], ''The Bank of Finland's Consumption Price Index'', *Bank of Finland Monthly Bulletin*, Vol. 10, pp.1-8.

COMMENTS

R. Robert Russell
Department of Economics
New York University

My professional life-long fear has been that one day Erwin Diewert would organize a conference and make me discuss a paper by Dale Jorgenson. It's a little like being asked to give a 10-minute critique of the *Encyclopedia of Economics*.

In this paper, Jorgenson and Slesnick review consumer theory – from integrability conditions to duality, the theory and application of consumer equivalence scales, the theory of aggregation, the econometrics of pooled time-series and cross-section data, the theory of cost-of-living indexes, and the theory of social choice and welfare economics – from Arrow's impossibility theorem right up to recent results on interpersonal comparability and cardinal utility. Another tour de force.

But the paper is much more than review; it is an invaluable contribution. Especially noteworthy is the approach to the econometric estimation of market demand functions obtained by aggregating demand functions of individuals with different tastes, incorporating the restrictions implied by the theory of individual consumer behaviour. As this approach has been reported before,[1] I concentrate my remarks on the other important contribution: the theory and application of social cost-of-living indexes derived from social welfare functions and individual demand functions. This is the first empirical application (of which I am aware) of the theory of social cost-of-living indexes.

To put the Jorgenson-Slesnick (J-S) approach in perspective, let me first set up the social cost-of-living-index problem in the manner of Pollak [1981]. The expenditure function of the r^{th} individual, E^r, is defined by

$$E^r(u_r,P) = \min_{X^r}\left\{P\cdot X^r \mid U^r(X^r) \geq u_r\right\}, \tag{1}$$

where u_r, X^r, and U^r are the utility level, consumption bundle, and utility function, respectively, of the r^{th} consumer, and P is the price vector. The aggregate expenditure function

E, conditional on the utility profile, $u = u_1,\dots,u_R$, is defined by

$$E(u_1,\dots,u_R,P) = \sum_r E^r(u_r,P). \tag{2}$$

For a fixed aggregate expenditure, M, and a fixed P, the equality,

$$E(u_1,\dots,u_R,P) = M, \tag{3}$$

defines a utility possibility frontier, labelled UPF(P,M). The utility possibility frontier represents the maximal utility for each individual given the utilities of all other individuals (and given M and P). It is illustrated for the case of two individuals in Figure 1.

The utility possibility frontier presumes individual optimization (i.e., efficiency) only; no elements of cardinal utility or interpersonal utility comparisons are involved. It shifts upward with increases in M and downward with increases in P; it is the social analogue of the individual budget constraint.

Social optimality – the choice of a point on the utility possibility frontier – requires the specification of a social welfare function, W. For a particular value of social welfare, ω, the equality,

$$W(u) = W(u_1,\dots,u_R) = \omega, \tag{4}$$

defines a social indifference curve (SIC(ω)). Social optimality (for given M and P) is represented by the tangency of a SIC(ω) and the UPF(P,M) (at $\overset{*}{u}$ in Figure 1).

The social expenditure function, ξ, is defined by

$$\xi(\omega,P) = \min_{u_1,\dots,u_R} \left\{E(u_1,\dots,u_R,P) \mid W(u) \geq \omega\right\}. \tag{5}$$

The Pollak social cost-of-living index, applied by Jorgenson and Slesnick, is

Figure 1

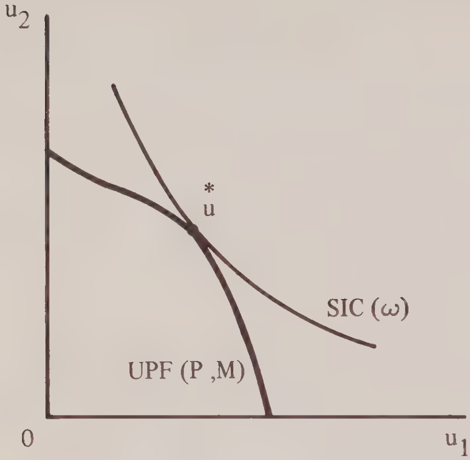


Figure 2

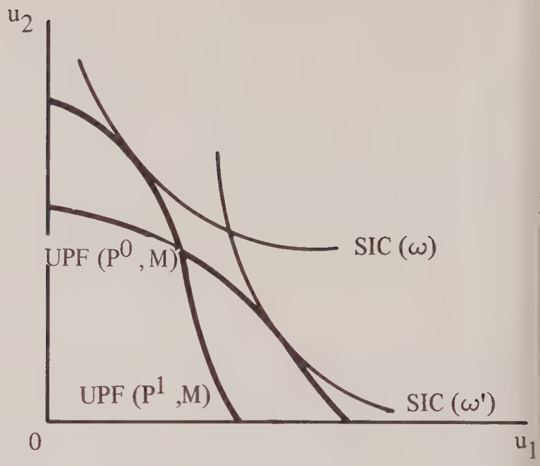
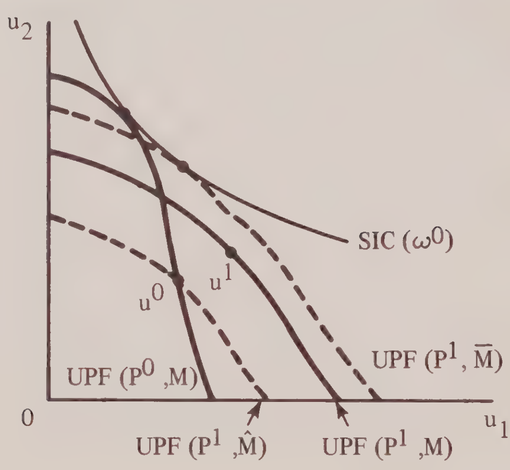


Figure 3



$$SCOL(\omega,P^0,P^1) = \frac{\xi(\omega,P^1)}{\xi(\omega,P^0)} . \qquad (6)$$

There are two fundamental problems with this approach. To illustrate the first, consider a change in prices from P^0 to P^1, holding M constant, with corresponding utility possibility frontiers reflected in Figure 2. Consider alternative social welfare functions, W and W', with social indifference curves labelled SIC(ω) and SIC (ω'), respectively. Given the social welfare function, W, the cost of living is lower in situation 0 than in situation 1, since M would have to be increased to raise UPF(P^0,M) to a point of tangency with SIC(ω). Given the social welfare function W', the cost of living is **higher** in situation 0 than in situation 1, since M must be increased to obtain a tangency of UPF(P^1,M) and SIC(ω').

Thus, diametrically opposite conclusions about the direction of change of the cost of living (and hence about changes in social welfare) result from the adoption of alternative social welfare functions. Of course, the same type of problem exists for an individual cost-of-living index: opposite conclusions about the direction of change of the cost of living can be obtained by the specification of alternative utility functions. But there is a crucial difference: assumptions about individual preferences are empirically testable, whereas assumptions about social welfare functions are not.

Jorgenson and Slesnick have jumped precariously from the familiar realm of positive econometrics to the ethereal realm of normative econometrics. The question is: who is to choose the social welfare function? I suspect that all would agree that neither a university economics professor nor a Bureau of Labor Statistics or Statistics Canada statistician should be saddled with this responsibility.

To be sure, Jorgenson and Slesnick are poignantly aware of this problem. It is for this reason that they painstakingly derive their social welfare function (using some powerful results of Roberts [1980]) from a set of axioms about the measurement of utility, interpersonal comparability, and social values (horizontal and vertical equity).[2] Nevertheless, this derivation – however ingenious – simply crystalizes the issue: none of these axioms is empirically refutable.[3] Moreover, the assumed social values, while couched euphemistically

in terms of equity, are in fact strongly egalitarian: social optimality requires equal incomes for "scale-equivalent" households.

Even if we are prepared to accept a particular social welfare function – say, that proposed by Jorgenson and Slesnick – there is a second fundamental problem. While the assumption of individual utility maximization is almost tautological so long as we are willing to accept consistency of individual behaviour, there can be no comparable presumption that social welfare is maximized. Indeed, since maximization of the J-S social welfare function requires equal incomes for (scale-equivalent) households, it is evident that society does **not** maximize this social welfare function.

The implication of this problem for the interpretation of cost-of-living indexes is illustrated in Figure 3. Shown in this figure are two utility possibility frontiers, denoted UPF(P^0,M) and UPF(P^1,M), reflecting efficient utility profiles in two different periods with the same aggregate income, M, but different prices, P^0 and P^1. Maximum social welfare in period 0 is ω^0, as indicated by the tangency of SIC(ω^0) and UPF(P^0,M).

The Pollak-Jorgenson-Slesnick social cost-of-living index would indicate an increase in the cost of living between periods 0 and 1, since aggregate income would have to be increased from M to $\bar{M}$ to attain social welfare ω^0 at prices P^1. That is, UPF(P^1,$\bar{M}$), which is tangent to the base-period social indifference curve, is everywhere above UPF(P^1,M) and hence $\bar{M}$ must be above M. Thus, the social cost-of-living index calculation would suggest that welfare was higher in period 0 than in period 1. But, if society does not in fact maximize social welfare, then this need not be the case. For example, if the actual utility profiles were given by u^0 and u^1 in Figure 3, then welfare in period 1 would be higher than in period 0 for any welfare function incorporating the Pareto principle. Since society does not in fact maximize social welfare, the nexus between changes in the cost of living at given income levels and changes in welfare, stressed so much by Jorgenson and Slesnick, is broken.

The J-S model is illustrated, for the case of two households, in Figure 4. The translog individual utility functions generate concave utility possibility frontiers. (This is easily shown by substituting the translog indirect utility functions into equation (2) and applying the

implicit function theorem to equation (3) to evaluate the rate of change u_i with respect to u_j). Along the ray where u_1 equals u_2, the slope of the utility possibility frontier is equal to the negative of the ratio of the scale factors, $-m_0(P,A_1)/m_0(P,A_2)$.[4]

The J-S social welfare function is a weighted average of individual-household indirect utility functions, with weights given by the household scale factors (normalized to sum to unity), plus a function of the scale factors and individual utilities that vanishes when all utilities are equal (a concave penalty function for income inequality).[5] The social indifference curves are therefore convex to the origin. Moreover, along the equal-utilities ray, the slopes of the indifference curves are equal to the negative of the ratio of the household scale factors, $-m_0(P,A_1)/m_0(P,A_2)$. (Were it not for the loss function for income inequality, each indifference curve would be linear with this slope.)

In Figure 4, $UPF(P^0,M^0)$ is the utility possibility frontier in situation 0. Social optimality in the J-S model requires that individual utilities be equated, as reflected by the tangency of the social indifference curve, $SIC(W(\overset{*}{u},P^0))$ and the utility possibility frontier at utility profile $\overset{*}{u}$. In reality, utilities are not equated and might be represented by a point such as u^0, with associated welfare $W(u^0,P^0)$.

Suppose that the utility possibility frontier shifts to $UPF(P^1,M^1)$ in situation 1. The question that underlies the J-S social cost-of-living index is this: What level of aggregate expenditure allows society to obtain the **optimal** social welfare in situation 1, $W(\overset{*}{u},P^0)$, at situation-2 prices, P^2? The answer is M^1, the aggregate income that generates the utility possibility frontier that contains $\overset{*}{u}$ (the situation-1 optimal utility profile) – vis., $UPF(P^1, M^1)$. The social welfare associated with the social indifference curve tangent to this utility possibility frontier, $W(\overset{*}{u},P^1)$, must be equal to the optimal social welfare in period 1, $W(\overset{*}{u},P^0)$. This is so even though the change in prices from P^0 to P^1 changes the weights in the social welfare function (and hence generates a new set of social indifference curves) because at equal incomes, such as at $\overset{*}{u}$, the social welfare is independent of the weights (and, of course, the penalty term for income inequality vanishes). Thus, in the J-S model, the social cost of living in situation 1 relative to situation 0 is equal to M^1/M_0, the ratio of aggregate income needed to obtain maximal social welfare in situation 0 at situation-1

Figure 4

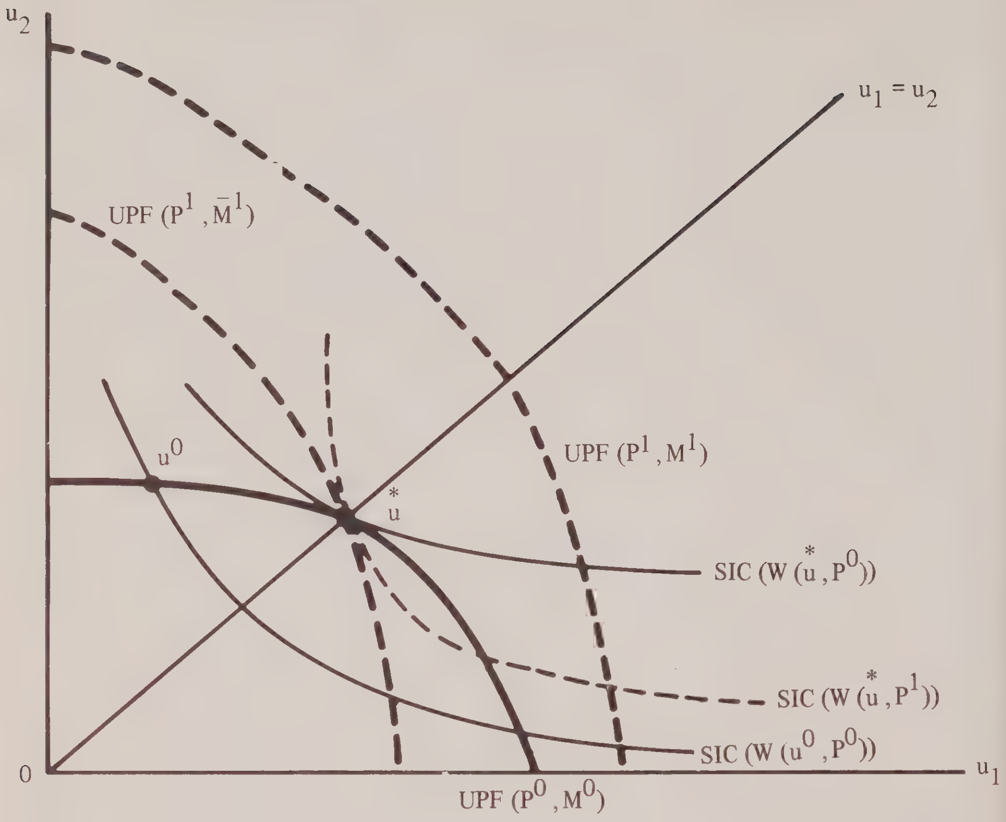
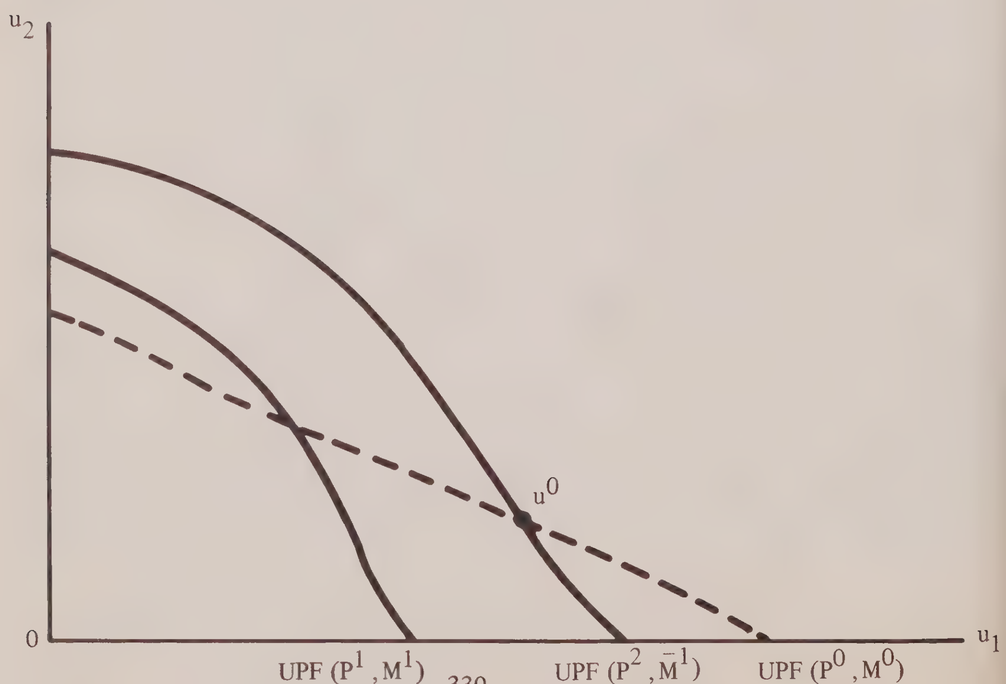


Figure 5



prices divided by actual aggregate income in situation 0. Thus, the J-S calculation compares only socially optimal points, even though society clearly does not optimize the J-S social welfare function.[6]

Thus, the J-S application of the theory of social cost-of-living indexes is plagued by both of the problems noted above. Cost-of-living comparisons are made along an equal-utilities ray (1) because of appealing but arbitrary normative assumptions and (2) despite the fact that actual utility profiles of society are not on this ray.

The question, of course, is this: What are the alternatives? Surely, as long as cost-of-living indexes are used for indexation of incomes of groups (such as unions, social security beneficiaries and welfare recipients), it makes little sense – at least from a theoretical point of view – to pretend that individual cost-of-living indexes are applicable for this purpose.[7] Three alternatives come to mind.

One alternative is to constrain individual utility functions sufficiently to ensure that the social cost-of-living index is independent of the form of the social welfare function. As Pollak [1981] showed, a sufficient condition for this independence is that aggregate demands be independent of the distribution of income. This is equivalent to individual preferences satisfying the condition of the Gorman polar form which is reflected in the following form of the expenditure function:

$$E^r(y,P) = \Lambda^r(P) + \pi(P)u_r ,$$

where Λ^r and π are positively linearly homogenous functions, and π is identical for all households.[8]

Equating the sum of these expenditure functions to aggregate expenditure (see equations (2)-(3)), we obtain the expression implicitly defining the utility possibility frontier:

$$\sum_r \Lambda^r(P) + \pi(P)\sum_r u_r = M.$$

It is apparent that the slope of the utility possibility frontier is everywhere equal to -1. Consequently the utility possibility frontiers generated by different price vectors and aggregate expenditures cannot cross, and the ranking of price-expenditure combinations is independent of the social welfare function. Thus, the cost-of-living calculation amounts simply to comparing utility possibility frontiers, with no meaningful reference to the social welfare function.

The trouble with this approach is that the assumption of linearity of the expenditure function in income is very strong. Moreover, empirical tests have rejected this property (see, for example, Pollak and Wales [1979]).

A second alternative is to eschew the social welfare approach altogether and to deal with cost-of-living changes for a group as a sort of statistical average. The idea is to calculate the change in the cost of living for an average (or representative) household with average income. The J-S model lends itself well to this approach, since the ''average person'' is well-defined in terms of the equivalence scales of their model. Indeed, where Jorgenson and Slesnick really end up, after invoking a medley of social choice results, is a statistical average. That is, the social cost-of-living function that they eventually employ is in fact the cost-of-living index for the average person with average income. Thus, one could just as well interpret the J-S calculations as those for the representative household, in which case one need not take into account their very strong assumptions about social values. The only assumptions that are relevant are those pertaining to individual utility functions, and these are empirically testable (at least in principle).

A third approach is to calculate a group cost-of-living index using utility possibility frontiers only – maintaining agnosticism about the existence as well as the form of a social welfare function. To illustrate, consider the base-period utility possibility frontier, $UPF(p^0,M^0)$, and base-period utility profile, u^0, in Figure 5. Suppose that the utility possibility frontier shifts to $UPF(p^1,M^1)$ in period 1. The proposed group cost-of-living index asks the following question: What aggregate income is required, at situation-1 prices, in order for the base-period utility profile u^0 to be feasible? In the picture, this expenditure level is $\overline{M}^1$, which, in combination with period-1 prices, P^1, generates a utility

possibility frontier that contains u^0. Formally, this cost-of-living index, which we might call the Laspeyres group cost-of-living index, is given by

$$GCOL(u^0,M^0,P^1) = \sum_r E^r(u_r^0,P^1)/M^0.$$

(Note that this index implicity depends on base-period prices and income distribution, since u^0 can be replaced by $V^1(P^0/M_1^0),\ldots,V^R(P^0/M_R^0)$.)

We could similarly define a Paasche group cost-of-living index by adopting the period-1 utility profile u^1 as the base – i.e., by asking how much the base-period aggregate income would have to be adjusted in order to allow each household to obtain the utility level of situation 1.

The advantage of this approach is that it deals only with empirically refutable information about individual preferences, prices, and incomes; no assumptions or information about cardinal utility, interpersonal utility comparisons, or social values are required. Of course, the disadvantage is that the approach cannot be linked to social welfare. But there is no way to avoid this defect unless one is willing to make the strong assumptions that Jorgenson and Slesnick need for their approach. And even then, as noted above, the link between the index and social welfare is broken if society does not maximize welfare.

This approach to group cost-of-living indexes is analogous to the theory of Laspeyres and Paasche indexes as applied to individual consumers. The utility possibility frontier is analogous to the budget constraint and the utility profiles are analogous to consumption bundles.

If one were willing to assume that society does, indeed, maximize some type of social welfare function, but the index-number analyst does not know its form, then I suspect one could prove theorems relating the true but unknown group cost-of-living index to these Paasche and Laspeyres group indexes – that is, bounding theorems, analogous to those in Pollak [1971] and Diewert [1982]. (The non-linearity of the utility possibility frontiers would be problematic, but I don't believe that this would preclude getting some results.)

I do not mean to imply that any of these three alternatives is clearly preferable to the J-S approach. On the contrary, they are inchoate notions. The J-S contribution, on the other hand, is concrete. They might underestimate the effort required to implement – and, more important, to update – their model, but Jorgenson and Slesnick have gone far beyond all previous studies in making the social cost-of-living index a practical concept. Undoubtedly, much additional work needs to be done, but Jorgenson and Slesnick are to be commended for a promising and ambitious initiation of research on the empirical application of social cost-of-living indexes.

Footnotes

1 The most immediate precursor is Jorgenson, Lau, and Stoker [1982]. For an earlier contribution to the empirical aggregation of demand systems (without the stringent requirements of exact aggregation in the sense of Gorman [1953]), see Berndt, Darrough, and Diewert [1977].

2 Their specific functional form for the income-inequality part of the social welfare function – a mean-value function – additionally requires a strong additivity assumption.

3 The authors seem to argue that cardinality is an implication of their aggregation conditions. This can't be correct. Linearity of the indirect utility function in expenditure (equation (2.6)) is sufficient, but not necessary, for aggregability. In short, the aggregation properties of their model are preserved by arbitrary monotone transformations of the indirect utility function.

4 In the J-S model, the indirect utility functions are identical for all households except for household-specific scale factors, $m_0(P,A_r)$, that enter as deflators of money income. That is, the utility functions are identical in prices and scale-adjusted incomes.

5 Since the scale factors are functions of prices, the J-S social welfare function is conditional on prices – directly as well as through the indirect effect of prices on individual utilities – i.e., the image is W(u,P).

6 The authors are not altogether clear about this fact. Equations (6.2) to (6.4), taken together, seem to imply that the social cost-of-living index is $W(u^0,P^0)/M^0$, the ratio of aggregate expenditure needed to obtain the **actual** social welfare in situation 0 at situation-1 prices, P^1, to the actual expenditure in situation 0. Such a comparison, however, would confound changes in the cost of living with changes in the distribution of income, since it would be comparing a social optimum in one situation with a suboptimal distribution of income in another. That this is not what the authors intend is indicated by the discussion following equations (6.2) to (6.4) and the discussion of Diagram 2, which compares only social optima.

7 It should be noted that indexation of individual and group incomes is but one use of a consumer price index. Equally important is the use of such indexes to evaluate the macroeconomic performance of an economy. For this purpose, an index need not be based on the theory of individual utility maximization or social welfare. What is required is a reasonable measure of price inflation, and mechanistic indexes can serve tolerably well. For the purpose of evaluating macroeconomic performance, it would be useful to link a consumer price index to the welfare costs of inflation (i.e., to the induced-inefficiency costs of inflation attributable to the decreased information content of prices and the welfare losses associated with the attendant redistribution of income). But this welfare aspect of consumer price indices is quite different from the welfare issues addressed by Jorgenson and Slesnick (and others in session).

8 The properties of the Gorman polar form are described in detail in Blackorby, Boyce, and Russell [1978] and Blackorby, Primont, and Russell [1978, ch. 8].

References

Berndt, E.R., M.N. Darrough and W.E. Diewert [1977], ''Flexible Functional Forms and Expenditure Distributions: An Application to Canadian Consumer Demand Functions'', *International Economic Review* 18, pp.651-676.

Blackorby, C., R. Boyce and R.R. Russell]1978], ''Estimation of Demand Systems Generated by the Gorman Polar Form: A Generalization of the S-Branch Utility Tree'', *Econometrica* 46, pp.345-364.

_________ D. Primont and R.R. Russell [1978], *Duality Separability, and Functional Structure: Theory and Economic Applications*, New York: North-Holland.

Diewert, W.E. [1983], ''The Theory of the Cost-of-Living Index and the Measurment of Welfare Change'', *Price Level Measurement: Proceedings From a Conference Sponsored by Statistics Canada.*

Gorman, W.M. [1953], ''Community Preference Fields'', *Econometrica* 21, pp.63-80.

Jorgenson, D.W., L.J. Lau and T.M. Stoker [1980], ''Welfare Comparison Under Exact Aggregation'', *American Economic Review* 70, pp.268-272.

Pollak, R.A. [1971], ''The Theory of the Cost-of-Living Index'', Research Discussion Paper II, Office of Prices and Living Conditions, U.S. Bureau of Labor Statistics, Washington, D.C.

_________ [1981], ''The Social Cost-of-Living Index'', *Journal of Public Economics* 15, pp.311-336.

_________ and T.J. Wales [1979], ''Welfare Comparisons and Equivalent Scales'', *American Economic Review* 69, pp.216-21

Roberts, K.W.S. [1980], ''Interpersonal Comparability and Social Choice Theory'', *Review of Economic Studies* 47, pp.421-439.

LEISURE TIME AND THE MEASUREMENT OF ECONOMIC WELFARE

W. Craig Riddell
Department of Economics
University of British Columbia

SUMMARY

Most measures of economic welfare and the cost of living are based on consumer expenditure on goods and services. Such measures omit an important commodity – leisure time – which is purchased implicitly by not working. This paper discusses the issues involved in incorporating leisure time in measures of the standard of living, and empirically implements the theoretical discussion by calculating some measures of economic welfare for Canada which incorporate leisure.

In order to calculate true measures of economic welfare, knowledge of the preferences of individual consumers or households is required. Pencavel [1977] provides measures of welfare for the U.S. using the estimated parameters of the Stone-Geary utility function. The advantage of this procedure is that it allows for substitution among goods and leisure in response to changes in relative prices. The disadvantage is that it assumes a specific functional form for the consumer's utility function.

This paper employs index number theory to obtain measures of economic welfare in the joint commodity demand-labour supply context. Thus the measures obtained are not based on any assumed functional form for the utility function. However, the index numbers do not fully allow for substitution in response to changes in relative prices and are thus only approximations to the true measures of changes in the standard of living.

A related issue addressed in this paper is the interpretation of published real wage or real earnings indexes, obtained by deflating nominal wages or earnings by a price index. While these are related to economic welfare, neither is a legitimate measure of the standard of living. Earnings may increase because of an increase in hours worked at existing

hourly wage rates or because of an increase in the hourly wage with no change in hours of work. The two have very different implications for welfare change. Similarly real wage indexes are not satisfactory measures of welfare because they do not allow for leisure time. However, real wage and real earnings indexes which are legitimate measures of the standard of living can be constructed; the details are given in the paper.

RÉSUMÉ

La plupart des mesures du bien-être économique et du coût de la vie sont basées sur les dépenses des consommateurs en biens et services. De telles mesures laissent de côté un bien important – le temps de loisirs – qui est acheté implicitement en ne travaillant pas. Le présent article discute les problèmes que présente l'incorporation du temps de loisirs dans les mesures du niveau de vie, et applique empiriquement la discussion théorique en calculant quelques mesures du bien-être économique au Canada qui comprennent les loisirs.

Pour calculer les mesures véritables du bien-être économique, il faut connaître les préférences des consommateurs ou des ménages particuliers. Pencavel [1977] fournit des mesures du bien-être aux États-Unis en utilisant les paramètres estimés de la fonction d'utilité Stone-Geary. L'avantage de cette technique est de permettre des substitutions entre les biens et les loisirs en réaction à des changements des prix relatifs. Son désavantage est de postuler une forme fonctionnelle spécifique à la fonction d'utilité du consommateur.

On utilise la théorie des nombres indices pour obtenir des mesures du bien-être économique dans le double contexte demande de produits/offre de main-d'oeuvre. Ainsi, les mesures obtenues ne sont pas basées sur un quelconque postulat de forme fonctionnelle pour la fonction d'utilité. Cependant, les nombres indices ne permettent pas pleinement les substitutions en réaction à des changements des prix relatifs et ne sont pas conséquent que des approximations des mesures réelles des changements du niveau de vie.

Un problème connexe traité ici est l'interprétation des indices publiés des salaires réels ou des gains réels, obtenus en déflatant les salaires nominaux ou les gains nominaux par

un indice de prix. Bien que ceux-ci soient reliés au bien-être économique, ni l'un ni l'autre ne constitue une mesure légitime du niveau de vie. Les gains peuvent augmenter à cause d'une augmentation du nombre d'heures de travail sans changement du salaire horaire ou à cause d'une augmentation du salaire horaire sans changement du nombre d'heures de travail. L'un ou l'autre ont des conséquences très différentes sur les changements de bien-être. De même, les indices de salaire réels ne constituent pas des mesures satisfaisantes du bien-être parce qu'ils n'incluent pas les loisirs. Cependant, il est possible de construire des indices de salaire réels et de gains réels qui soient des mesures légitimes du niveau de vie; le texte en fournit les détails.

Il est possible de baser des mesures du bien-être économique sur n'importe laquelle des variables qui sont exogènes aux décideurs. Dans la théorie classique du comportement du consommateur, le revenu et les prix des biens sont exogènes et, parce que des comparaisons basées sur le prix d'un bien particulier risquent d'avoir peu de signification, on utilise généralement le revenu comme mesure du bien-être. Dans le cas d'un comportement relié de consommation et d'offre de main-d'oeuvre, les variables exogènes sont les prix, le taux de salaire, le revenu de sources autres que le travail et le temps total disponible. Les mesures de bien-être basées sur le taux de salaire, sur le revenu provenant d'autres sources que le travail et "le plein revenu" (le revenu que l'on gagnerait si l'on passait tout le temps disponible à travailler) sont décrites dans ce document. Bien que toutes trois soient des mesures conceptuellement valides du changement du niveau de vie, certaines devraient être plus utiles que d'autres. Le désavantage des mesures du revenu réel provenant d'autres sources que le travail est que celui-ci ne constitue, pour la plupart des gens, qu'une composante relativement petite du revenu total. Les mesures de plein revenu ont le désavantage d'offrir peu d'intérêt intuitif (peu de gens pensent en termes de leur plein revenu). De plus, ce qui est peut-être plus important, les mesures de plein revenu contiennent un élément d'arbitraire dans la détermination du temps total disponible pour le travail et les loisirs. Par ailleurs, le revenu total est une extension naturelle du revenu ou de la dépense à la situation de consommation et d'offre de main-d'oeuvre conjoints. Ceci devient particulièrement clair si l'on généralise l'indice Tornqvist au modèle du choix revenu/loisir. De plus, les indices de plein revenu établissent des frontières plus strictes aux mesures véritables de bien-être que ne le font les mesures du revenu de sources autres que du travail (Théorème 6).

Les mesures de bien-être basées sur le taux de salaire ont un intérêt intuitif raisonnable, ne sont arbitraires en aucune façon et présentent l'avantage que le revenu du travail est la principale composante du revenu pour la plupart des gens. On discute les différences qui existent entre ces mesures et les indices classiques des salaires réels.

Les nombres indices des taux de salaire constituent une alternative intéressante à ceux qui sont basés sur le revenu provenant d'autres sources que le travail ou le plein revenu. Une autre possibilité, suggérée récemment par Diewert [1982], est un indice basé sur les dépenses de consommation au cours de la période de base. C'est là un indice du type Malmquist, qui présente deux avantages: (1) il est basé sur une quantité qui est intuitivement intéressante, les dépenses en biens et (2) il n'exige pas d'attribution arbitraire du temps total disponible.

Ces divers nombres-indices du bien-être économique sont calculés pour le Canada à l'aide des données annuelles pour la période allant de 1949 à 1980. Les indices de taux de salaire indiquent que les salaires réels ont augmenté plus qu'on pourrait l'estimer sur la base des mesures traditionnelles de "salaire réel" qui ne tiennent pas compte du choix revenu/loisir. De même, les indices de dépenses de consommation de Diewert indiquent que le bien-être économique a augmenté d'environ 240% au cours de la période 1949-1980 plutôt que de 200% qu'on obtient avec les indices classiques de "revenu réel".

1. Introduction

One of the most important commodities (indeed, probably **the** most important commodity omitted from measures of economic welfare and the cost of living is leisure time. There is, of course, a good reason for this omission. Leisure time is purchased implicitly (by not working) rather than explicitly, and the price associated with an hour of leisure is also implicit. Nonetheless there are also good reasons for attempting to incorporate leisure into measures of the cost of living, real wages and real income. Observed changes in earnings may have come about because of changes in wage rates or changes in hours worked or various combinations of the two. Yet for measures of welfare the composition of the change in earnings is important. Further, observed changes in earnings include hours of work decisions which themselves reflect adjustments to changes in wage rates and these underlying time allocation choices should be incorporated in measures of welfare.

This paper will discuss the issues involved in incorporating leisure time in measures of the cost of living, real wages and real income and will empirically implement the discussion by deriving some index numbers which incorporate leisure.

The next section of the paper briefly summarizes the theory of index numbers of real income in the context of the standard model of consumer behaviour. This material is well-known, and is included here because it provides useful background for the theory of index numbers of real wages or real income in the context of the income-leisure choice model of consumer-worker behaviour. Section 3 contains this discussion of the theory of index numbers of real income in the joint consumption and labour supply case. This material is not well-known, the main contributions having been made quite recently (specifically, Pencavel [1977, 1979a], Lloyd [1979], Cleeton [1982]).

The CPI is widely used as a measure of the cost of living. It is a relatively straightforward matter to incorporate leisure time into measures of the cost of living, if this is desired. The Family Expenditure Survey, upon which the weights of CPI are based, could be expanded to include questions about hours of work and hourly wage rates. Hours of leisure could be treated in the same fashion as an explicitly purchased commodity, the price of leisure being the after-tax hourly wage rate. The main complications are two: (1) the marginal price (the price of an additional hour of leisure) may not equal the average price, because of tax considerations (with a progressive income tax the marginal after-tax wage may differ from the average after-tax wage) and possibly because of hours of work constraints (increasing leisure time may involve switching jobs, at least in the short run; reducing leisure may involve a second job with a different wage than the primary job), and (2) the price of leisure differs across individuals (because of differences in after-tax wage rates) to a much greater extent than do the prices of other commodities. Nonetheless, these complicating factors are not particularly severe.

While some discussion of measures of the cost of living will be provided, the main focus of the paper is on measures of the standard of living, real income, or economic welfare (these three terms will be used synonomously). The CPI is frequently used to deflate money incomes or money wages, producing measures of real incomes or real wages. Are these legitimate measures of the standard of living? Real income measures are typically based on the standard theory of consumer behaviour in which the consumer's money income

is taken to be exogenous. However, once the individual's time constraint and the concomitant labour supply decision are recognized, the income available for expenditure on goods is chosen by the consumer-worker. In this setting, most measures of real income are not true measures of the standard of living. Section 3 of the paper discusses two concepts that can be used for this purpose: real "non-labour income" and real "full income" (defined later).

Similarly, deflating nominal wages by the CPI does not produce a true measure of economic welfare. Real wage indices that have this property can, however, be constructed and these are discussed in Section 3. Because of the importance of labour income to most individuals, such real wage measures are likely to be of more interest than real non-labour income measures.

Pencavel [1977] appears to have been the first author to attempt to construct index numbers of real wages and real non-labour income that are true measures of the standard of living. His procedure employed the estimated parameters of a direct utility function and was based on the income-leisure choice model. (The estimates were obtained from Abbott and Ashenfelter's [1976] study of commodity demand and labour supply.) The advantage of this procedure is that it allows for substitution among goods and leisure in response to changes in relative prices. The disadvantage is that it assumes a specific functional form for the consumer's utility function (the Stone-Geary utility function in Pencavel's study).

This paper employs index number theory to obtain measures of real income and real wages in the joint commodity demand-labour supply context. Thus the measures obtained are not based on any assumed functional form for the utility function. However, the index numbers do not fully allow for substitution in response to changes in relative prices and are thus only approximations to the true measures of changes in the standard of living.

2. Index Numbers of Real Income and The Cost of Living

The purpose of this section is to summarize some known results on index numbers which are needed later in the paper. Proofs of these results together with references to the previous literature may be found in the survey paper of Diewert [1981].

Measures of real income and the cost of living are usually based on the standard model of consumer behaviour. Let U(x) be the individual's ordinal utility function where $x = (x^1 \ldots x^n)$ is a 1 by n vector and x^i is the quantity of the ith good. Let $p = (p^1 \ldots p^n)$ be the vector of prices of the n goods and y be the consumer's income. If p and y are exogenous, the consumer will choose quantities consumed to

$$\max_{(x)} U(x) \qquad \text{(A)}$$

$$\text{subject to } p \cdot x = y$$

where $p \cdot x$ is the inner product of the vectors p and x.

Let $x^*(p,y)$ be the solution to (A); this is the set of commodity demand functions. Substituting into U(x) gives the indirect utility function

$$V(p,y) = U(x^*(p,y)) \qquad (1)$$

which shows the maximum utility as a function of the exogenous variables prices and income. Inverting (1) gives the expenditure function $e(p,\bar{U})$ which solves

$$\min_{(x)} p \cdot x \qquad \text{(B)}$$

$$\text{subject to } U = \bar{U}$$

That is, $e(p,\bar{U})$ shows the minimum expenditure needed to achieve utility $\bar{U}$ at prices p.

Now suppose that prices and income vary over time (or across regions or countries) and we wish to construct measures of the consumer's welfare or standard of living. Let (p_0,x_0,y_0) and (p_1,x_1,y_1) be the observed prices, quantities consumed and expenditure in periods 0 and 1 respectively and let $Q(p_0,p_1,x_0,x_1)$ be an index number of real income. Such an index should have the property

$$Q(p_0,p_1,x_0,x_1) > 1 \text{ if and only if } V(p_1,y_1) > V(p_0,y_0) \tag{P1}$$

That is, the index ''must itself be a cardinal indicator of ordinal utility'' (Samuelson and Swamy, [1974, p.568]). Two such measures are

$$Q_{LA} \equiv e(p_0,U(x_1))/e(p_0,U(x_0)) \tag{2}$$

$$Q_{PA} \equiv e(p_1,U(x_1))/e(p_1,U(x_0)) \tag{3}$$

Q_{LA} is the Laspeyres-Allen index of real income and Q_{PA} is the Paasche-Allen index of real income. These true, or constant-utility, index numbers of real income are, of course, unobservable. However, $p_0 \cdot x_1 \geq e(p_0,U(x_1))$ and $p_1 \cdot x_0 \geq e(p_1,U(x_0))$ so that upper and lower bounds can be obtained. Thus

$$Q_{LA} \leq \frac{p_0 \cdot x_1}{p_0 \cdot x_0} \equiv Q_L \tag{4}$$

$$Q_{PA} \geq \frac{p_1 \cdot x_1}{p_1 \cdot x_0} \equiv Q_P \tag{5}$$

where Q_L is the Laspeyres quantity index and Q_P is the Paasche quantity index.[1]

Another constant-utility index number of real income is the Malmquist index. The Laspeyres-Malmquist index Q_{LM} and Paasche-Malmquist index Q_{PM} are given by

$$Q_{LM} \equiv D(U(x_0), x_1) \tag{6}$$

$$Q_{PM} \equiv 1/D(U(x_1), x_0) \tag{7}$$

where $D(U(x), \bar{x})$ is the deflation function defined by

$$D(U,\bar{x}) \equiv \max_{k>0} \left\{k : U(\bar{x}/k) \geq U\right\} \tag{8}$$

One advantage of the Malmquist indices is that both upper and lower bounds exist for each:

$$\text{Min}\left\{x_1^i/x_0^i;\ i=1,\ldots.n\right\} \leq Q_{LM} \leq Q_L \tag{9}$$

and

$$Q_P \leq Q_{PM} \leq \underset{i}{\text{Max}}\left\{x_1^i/x_0^i;\ i=1\ldots.n\right\} \tag{10}$$

Two other quantity indices are Fisher's ideal quantity index Q_F and the Tornqvist or translog quantity index Q_T. These are defined by

$$Q_F \equiv (Q_L \cdot Q_P)^{1/2} \tag{11}$$

$$Q_T \equiv \prod_{i=1}^{n} (x_1^i/x_o^i)^{1/2(s_o^i\ +\ s_1^i)} \tag{12}$$

where s_t^i is commodity i's share of total expenditure in period t.

$$s_t^i \equiv p_t^i x_t^i / p_t \cdot x_t \tag{13}$$

Turning now to index numbers of the cost of living, the Laspeyres-Konüs P_{LK} and Paasche-Konüs P_{PK} constant-utility index numbers are

$$P_{LK} \equiv e(p_1,U(x_0))/e(p_0,U(x_0)) \tag{14}$$

$$P_{PK} \equiv e(p_1,U(x_1))/e(p_0,U(x_1)) \tag{15}$$

These are bounded by

$$\underset{i}{\text{Min}}\left\{p_1^i/p_0^i;\ i=1,\ldots,n\right\} \leq P_{LK} \leq P_L \tag{16}$$

$$P_P \leq P_{PK} \leq \underset{i}{\mathrm{Max}} \left\{p_1^i/p_0^i;\ i=1,\ldots,n\right\} \tag{17}$$

where P_L and P_P are the Laspeyres and Paasche price indices defined by

$$P_L \equiv p_1 \cdot x_0 / p_0 \cdot x_0 \tag{18}$$

$$P_P \equiv p_1 \cdot x_1 / p_0 \cdot x_1 \tag{19}$$

The CPI is, of course, a Laspeyres price index.

Two additional price indices are Fisher's ideal price index P_F and the Tornqvist or translog price index P_T:

$$P_F \equiv (P_P P_L)^{1/2} \tag{20}$$

$$P_T \equiv \prod_{i=1}^{n} (p_1^i/p_0^i)^{1/2(s_0^i + s_1^i)} \tag{21}$$

The standard model of consumer behaviour, and thus the index numbers based on that model, have a number of important limitations. Most important for present purposes is that while prices are typically exogenous to the individual consumer or household, income is not. Income and expenditure on goods are chosen by individuals. Three key aspects of this choice are the following. First, individuals facing exogenous hourly wages can determine income by choosing the number of hours per week and year to work. The ultimate constraint the individual faces is the time, not income, constraint. Second, individuals can influence the (hourly) wage rate they receive by investing in human capital (education, training, migration, etc.). Third, expenditure on goods need not equal income each period.

The latter two aspects influencing observed income and expenditure require an intertemporal choice model. Incorporating such life cycle aspects is outside of the objectives of this paper. The first aspect involves incorporating leisure as a commodity chosen by the individual. The most basic model which accomplishes this is the income-leisure choice model, and it is to the measurement of the standard of living in the context of this model that we now turn.

3. Income-Leisure Choice and Index Numbers of Real Income

The individual now chooses the vector of goods x and hours of leisure l to

$$\max_{(x,l)} U(x,l) \tag{B}$$

subject to $\quad p \cdot x = wh + m$

and $\quad h + l = T.$

Here h is hours worked, T is the total time available for the two activities, work and leisure, m is non-labour income, and w is the after-tax hourly wage rate. The variables T, m and w are taken to be exogenous, although as mentioned above both m and w would become choice variables in an intertemporal model. The price of leisure is the wage rate, as is made clear by writing the constraint as follows:

$$p \cdot x + wl = wT + m = Y \tag{22}$$

where $Y = wT + m$ is called "full income" and is exogenous to the consumer.[2] The constraint (22) is equivalent to that in the maximization problem (A); the left-hand side is total expenditure on goods including leisure and the right-hand side is exogenous income.

Measures of welfare can be based on any of the variables which are exogenous to the decision maker. In the standard theory of consumer behaviour, income and prices are treated as exogenous. Either could be used to construct constant utility index numbers; income is usually chosen because comparisons based on income are more meaningful than those based on any particular price. In the case of joint consumption and labour supply behaviour, the exogenous variables are prices p, the wage rate w, non-labour income m and total time available. Again, comparisons based on the price of a particular market good are unlikely to be very meaningful. Comparisons based on the price of the commodity leisure are, however, meaningful. Thus index numbers of real wages are potentially useful as measures of welfare change. Similarly, index numbers of real non-labour income may be employed.

Finally, index numbers of time available could also be constructed, although I am not aware of any such measures having been calculated. Presumably this reflects the fact that total time available T is not only exogenous but also fixed. Index numbers based on T would involve the conceptual experiment of altering the time available to the consumer-worker.

Attention can thus be focused on two welfare measures – index numbers of real wages and real non-labour income. As before, these are defined in terms of the value of the exogenous variables in the current period that would provide the consumer-worker with the same utility as in the base period.

With joint commodity demand and labour supply decisions, the indirect utility function becomes $V(p,w,m)$ and the expenditure function becomes $c(p,w,U)$. (The argument T has been suppressed because T is fixed.) A constant-utility index of real non-labour income is defined analogously to that of real income in the previous section. Define m_1^* as the minimum non-labour income required in period 1 to attain base period utility $V(p_0,w_0,m_0)$; that is;

$$m_1^* = c(p_1,w_1,U(x_0,l_0)) \tag{23}$$

Thus

$$V(p_1,w_1,m_1^*) = V(p_0,w_0,m_0) \tag{24}$$

and m_t/m_t^* is a constant-utility index number of real non-labour income. m_1/m_1^* corresponds to the Laspeyres-Allen index Q_{LA} defined earlier in the context of the standard theory of consumer behaviour. Similarly there is a Paasche-Allen index number of real non-labour income which uses period 1 prices and wage rate as the base. These two measures are given by

$$M_{LA} \equiv c(p_0,w_0,U(x_1,l_1))/c(p_0,w_0,U(x_0,l_0)) \tag{25}$$

$$M_{PA} \equiv c(p_1,w_1,U(x_1,l_1))/c(p_1,w_1,U(x_0,l_0)) \tag{26}$$

Note that M_{LA} and M_{PA} may be negative. Indeed, with wage rates rising over time we would expect $c(p_0,w_0,U(x_1,l_1))$ to become large relative to $c(p_0,w_0,U(x_0,l_0))$ so that M_{LA} will become large and we would expect $c(p_1,w_1,U(x_0,l_0))$ to possibly become negative, making M_{PA} negative.

Just as the Laspeyres-Allen index of real income Q_{LA} is bounded from above by the Laspeyres quantity index Q_L and the Paasche-Allen index of real income Q_{PA} is bounded from below by the Paasche quantity Q_P, the index numbers of real non-labour income are bounded by corresponding quantity indices M_L and M_P defined by

$$M_L \equiv (p_0 \cdot x_1 - w_0 h_1)/(p_0 \cdot x_0 - w_0 h_0) \tag{27}$$

$$M_P \equiv (p_1 \cdot x_1 - w_1 h_1)/(p_1 \cdot x_0 - w_1 h_0) \tag{28}$$

Theorem 1 (Cleeton [1982, p.224]):

$$M_{LA} \leq M_L \tag{29}$$

$$M_{PA} \geq M_P \tag{30}$$

Proof: From utility maximization, the denominators of M_{LA} and M_L are equal; that is,

$$p_0 \cdot x_0 - w_0 h_0 = m_0 = c(p_0,w_0,U(x_0,l_0)) \tag{31}$$

Comparing the numerators,

$$\begin{aligned} c(p_0,w_0,U(x_1,l_1)) &= \text{Min}\left\{(p_0 \cdot x - w_0 l) : U(x,l) \geq U(x_1,l_1)\right\} \\ &\leq p_0 x_1 - w_0 h_1 \end{aligned} \tag{32}$$

since (x_1,l_1) is feasible for the minimization problem but not necessarily optimal. The proof of (30) is similar.

Index numbers could be defined in terms of full income Y rather than non-labour income m. Let f(p,w,U) be the minimum full income expenditure function; that is,

$$f(p,w,\bar{U}) \equiv \underset{(x,l)}{\text{Min}} \left\{(p \cdot x + wl) : U(x,l) \geq \bar{U}\right\} \tag{33}$$

The Allen index numbers of real full income are

$$Y_{LA} \equiv f(p_0,w_0,U(x_1,l_1))/f(p_0,w_0,U(x_0,l_0)) \tag{34}$$

$$Y_{PA} \equiv f(p_1,w_1,U(x_1,l_1))/f(p_1,w_1,U(x_0,l_0)) \tag{35}$$

and the Laspeyres and Paasche quantity index numbers are

$$Y_L \equiv (p_0 \cdot x_1 + w_0 l_1)/(p_0 \cdot x_0 + w_0 l_0) \tag{36}$$

$$Y_P \equiv (p_1 \cdot x_1 + w_1 l_1)/(p_1 \cdot x_0 + w_1 l_0) \tag{37}$$

These quantity index numbers provide one-sided bounds for the corresponding constant-utility index numbers of full income.

Theorem 2:

$$Y_{LA} \leq Y_L \tag{38}$$

$$Y_{PA} \geq Y_P \tag{39}$$

Proof: Similar to Theorem 1.

Index numbers of real non-labour income are more intuitively appealing than those of real full income as few, if any, individuals think in terms of their full income (i.e., the sum of non-labour income and what labour income would be if no leisure time were consumed). This suggests using the quantity index numbers M_L and M_p in lieu of Y_L and

Y_P. One point, however, which may be considered an argument in favour of the full income quantity indices is that they provide tighter bounds than their non-labour income counterparts.

Theorem 3:

$$Y_L - Y_{LA} \leq M_L - M_{LA} \tag{40}$$

$$Y_{PA} - Y_P \geq M_{PA} - M_P \tag{41}$$

Proof: From the definitions of the minimum non-labour income and minimum full income functions

$$c(p_0,w_0,U(x_1,l_1)) + w_0T = f(p_0,w_0,U(x_1,l_1)) \tag{42}$$

Using the definitions (25) and (27), utility maximization implies that

$$M_L - M_{LA} = (p_0{\cdot}x_1 - w_0h_1 - c(p_0,w_0,U(x_1,l_1)))/m_0 \tag{43}$$

and, using (34) and (36),

$$Y_L - Y_{LA} = (p_0{\cdot}x_1 + w_0l_1 - f(p_0,w_0,U(x_1,l_1)))/Y_0 \tag{44}$$

Now note that (42) implies that the numerator of the right-hand side of (43) equals the numerator of the right-hand side of (44). However, $m_0 \leq Y_0$ which establishes (40). The proof of (41) is similar.

While Theorem 3 provides some justification for index numbers of real full income, two factors make use of such measures unattractive. One was mentioned above – few individuals think in terms of full income, so the measure does not have broad intuitive appeal. Second, the indices Y_{LA} and Y_{PA} contain an element of arbitrariness in the measurement of leisure l. Do we take T = 24 hours per day or T = 17, allowing seven hours for

activities which are neither leisure nor paid work (mainly sleeping and eating). A variety of alternative assumptions have been made in empirical work; see the discussion in Pencavel [1979b] or Usher [1980, pp.135-47]. This element of arbitrariness is not present in the non-labour income indices M_{LA} and M_{PA}.

Thus measures of both real non-labour income and real full income have some unattractive features. Because of the importance of labour income to most individuals, an appealing alternative is a real wage index. Indeed, real wages are frequently used along with real income as indicators of economic welfare.

Pencavel [1977] appears to have been the first author to construct index numbers of real wages that are true measures of the standard of living. The theory of real wage indices for this purpose was subsequently extended by Cleeton [1982]. To begin, define the minimum wage function as

$$w(p,m,\bar{U}) \equiv \min_{(x,l)} \left\{(p \cdot x - m)/h \; ; \; U(x,l) > \bar{U}, h + l = T\right\} \tag{45}$$

Thus $w(p,m,\bar{U})$ shows the minimum hourly wage rate required to attain utility $\bar{U}$ at prices p and non-labour income m. $w(p,m,\bar{U})$ is not defined if the individual's non-labour income is sufficiently great that $\bar{U}$ can be attained without working; i.e., if $U(x^*,T) > \bar{U}$ where $p \cdot x^* \leq m$.

The Allen index numbers of real wages can be defined in an analogous fashion to real income indices:

$$W_{LA} \equiv w(p_0,m_0,U(x_1,l_1))/w(p_0,m_0,U(x_0,l_0)) \tag{46}$$

$$W_{PA} \equiv w(p_1,m_1,U(x_1,l_1))/w(p_1,m_1,U(x_0,l_0)) \tag{47}$$

These true cost-of-living index numbers of real wages can be approximated by their Laspeyres and Paasche quantity index counterparts:

$$W_L \equiv [(p_0 \cdot x_1 - m_0)/h_1]/w_0 \tag{48}$$

$$W_P \equiv w_1/[p_1 \cdot x_0 - m_1)/h_0] \tag{49}$$

Theorem 4 (Cleeton [1982, p.223]):

$$W_{LA} \leq W_L \tag{50}$$

$$W_{PA} \geq W_P \tag{51}$$

Proof: From utility maximization,

$$w_0 = w(p_0, m_0, U(x_0, l_0)) \tag{52}$$

Comparing the numerators of (46) and (48):

$$\begin{aligned} W(p_0, m_0, U(x_1, l_1)) &= \mathrm{Min} \left\{ (p_0 \cdot x - m_0)/h : U(x,l) \geq U(x_1, l_1) \; ; h + l = T \right\} \\ &\leq (p_0 \cdot x_1 - m_0)/w_0 \end{aligned} \tag{53}$$

since (x_1, l_1) is feasible for the minimization problem but not necessarily optimal. The proof of (51) is similar.

Note that neither of the indices (46) or (47) or their observable counterparts (48) and (49) are "real wage indices" as the term is generally used. Deflating the money wage by the CPI does not provide a true measure of the standard of living for two reasons: (1) the contribution of non-labour income to welfare is ignored and (2) the contribution of leisure time to welfare is ignored.

The literature on index numbers of real wages and real income in the context of the income-leisure choice model has confined its attention to Allen-type indices of economic welfare. We now turn to the nature and properties of Malmquist-type indices in this setting.

With joint consumption and labour supply decisions, the Laspeyres-Malmquist index R_{LM} and Paasche-Malmquist index R_{PM} can be defined as follows:

$$R_{LM} \equiv D(U(x_0,l_0),(x_1,l_1)) \tag{54}$$

$$R_{PM} \equiv 1/D(U(x_1,l_1),(x_0,l_0)) \tag{55}$$

where the deflation function D is defined analogously to (8):

$$D(U,(\bar{x},\bar{l})) \equiv \text{Max}\left\{k : U(\bar{x}/k,\bar{l}/k) \geq U\right\} \tag{56}$$

Note that the Malmquist index of real income is not defined in terms of the amount of any particular exogenous variable required to attain the reference level of economic welfare. Thus there is no need to consider Malmquist indices of non-labour income, full income, or the wage rate. Nonetheless, any one of these three variables may be used to approximate a Malmquist true welfare index. For this reason it is useful to state results on the bounds of the Malmquist indices in two stages.

Theorem 5 (Diewert [1981, p. 175]):

$$\underset{i}{\text{Min}}\left\{x_1^i/x_0^i;\ i=1,....n+1\right\} < R_{LM} < \underset{i}{\text{Max}}\left\{x_1^i/x_0^i;\ i=1,....n+1\right\} \tag{57}$$

$$\underset{i}{\text{Min}}\left\{x_1^i/x_0^i;\ i=1,....n+1\right\} \leq R_{PM} \leq \underset{i}{\text{Max}}\left\{x_1^i/x_0^i;\ i=1,....n+1\right\} \tag{58}$$

$$\text{where} \quad x_t^{n+1} \equiv l_t \tag{59}$$

Proof: The theorem is a straightforward generalization of Theorem 13 in Diewert [1981]. Adding the labour supply or leisure choice dimension does not alter the structure of the proof.

The bounds given in (57) and (58) may not be particularly tight ones. In the standard consumer theory (income treated as exogenous) the Malmquist indices are also bounded

by the Laspeyres and Paasche quantity indices, and these provide a tighter upper bound for Q_{LM} and a tighter lower bound for Q_{PM} (recall equations (9) and (10)). In the income-leisure choice model we have discussed three alternative sets of Laspeyres and Paasche quantity indices (defined in terms of non-labour income, full income and the wage rate) and we now wish to investigate the relationship between these and the Malmquist indices R_{LM} and R_{PM}.

Theorem 6:

$$R_{LM} \leq Y_L \leq \underset{i}{\mathrm{Max}} \left\{x_1^i/x_0^i;\ i=1,\ldots n+1\right\} \tag{60}$$

$$\underset{i}{\mathrm{Min}} \left\{x_1^i/x_0^i;\ i=1,\ldots n+1\right\} \leq Y_P \leq R_{PM} \tag{61}$$

Proof: (a) Proof that $Y_L \leq \underset{i}{\mathrm{Max}} \left\{x_1^i/x_0^i\ ;\ i=1,\ldots n+1\right\}$

$$Y_L = \sum_{i=1}^{n+1} p_0^i x_1^i/(p_0 \cdot x_0 + w_0 l_0) \qquad \text{from (36) and (59)}$$

$$= \sum_{i=1}^{n} s_1^i x_1^i/x_0^i \qquad \text{from (13)}$$

$$\leq \underset{i}{\mathrm{Max}}\left\{x_1^i/x_0^i;\ i=1,\ldots n+1\right\}$$

since a share-weighted average must be less than the maximum value entering into the average.

(b) Proof that $R_{LM} \leq Y_L$

Let k^* be the solution to (54) and (56). That is,

$$R_{LM} \equiv \underset{k>0}{\text{Max}} \left\{ k : U(x_1/k, l_1/k) \geq U(x_0,l_0) \right\}$$

$$= k^*$$

Now $\quad p_0 \cdot x_0 + w_0 l_0 = f(p_0, w_0, U(x_0,l_0))$

$$\equiv \underset{x,l}{\text{Min}} \left\{ p_0 \cdot x + w_0 l : U(x,l) \geq U(x_0,l_0) \right\}$$

$$\leq p_0 \cdot x_1/k^* + w_0 l_1/k^*$$

since $(x_1/k^*, l_1/k^*)$ is feasible for the minimization problem.

Thus

$$k^* \leq (p_0 \cdot x_1 + w_0 l_1)/(p_0 \cdot x_0 + w_0 l_0) \equiv Y_L$$

The proof of (61) is similar

Theorem 6 provides additional justification for using index numbers of real full income rather than non-labour income. In particular, it is not true that

$$M_L \leq \underset{i}{\text{Max}} \left\{ x_1^i / x_0^i ;\ i = 1, \ldots n+1 \right\} .$$

The Tornqvist index generalizes easily to the income-leisure choice model.

$$Y_T \equiv \prod_{i=1}^{n+1} (x_1^i / x_0^i)^{1/2(s_0^i + s_1^i)} \tag{62}$$

where the shares are defined in terms of full income

$$Y = p \cdot x + wl = \sum_{i=1}^{n+1} p_i x_i \tag{63}$$

The Fisher ideal index can be defined in terms of full income

$$Y_F \equiv (Y_L\ Y_P)^{1/2} \tag{64}$$

However it is not possible to define the equivalent Fisher index in terms of non-labour income because M_L and M_P may be negative. The real wage indices do not have this property, however, so that we could define

$$W_F \equiv (W_L\ W_P)^{1/2} \tag{65}$$

Real wage index numbers are one alternative to index numbers based on non-labour income or full income. Another alternative, suggested recently by Diewert [1983] is an index based on consumption expenditure in the base period. Specifically, Diewert proposes the measures

$$R_{LDM} \equiv F(U(x_0,l_0),(x_1,l_1)) \tag{66}$$

$$R_{PDM} \equiv 1/F(U(x_1,l_1),(x_0,l_0)) \tag{67}$$

where the deflation function is defined in terms of consumption of commodities other than leisure:

$$F(U,(\bar{x},\bar{l}\ )) \equiv \text{Max}\left\{k\text{: } U(\bar{x}/k,\bar{l}\ ) \geq U\right\} \tag{68}$$

This deflation function can be compared to the function D in (56) which deflates the vector $(\bar{x}, \bar{l})$ rather than the goods consumption component $\bar{x}$. Geometrically, the Laspeyres-Diewert-Malmquist and Paasche-Diewert-Malmquist index numbers R_{LDM} and R_{PDM} deflate goods consumption along a vertical ray based on base period leisure time whereas the index numbers R_{LM} and R_{PM} deflate goods consumption and leisure time along a ray through the origin.

Observable approximations to R_{LDM} and R_{PDM} are the Laspeyres and Paasche consumption quantity index numbers E_L and E_P:

$$E_L \equiv \frac{p_0 \cdot x_1}{p_0 \cdot x_0 + w_0(h_1 - h_0)} \tag{69}$$

$$E_P \equiv \frac{p_1 \cdot x_1 + w_1(h_0 - h_1)}{p_1 \cdot x_0} \tag{70}$$

Theorem 7 (Diewert [1982, p.47]):

$$R_{LDM} \leq E_L \tag{71}$$

$$R_{PDM} \geq E_P \tag{72}$$

Proof: Similar to Theorem 1.

The index numbers E_L and E_P have two advantages over other measures which also incorporate leisure time: (1) they are based on an intuitively appealing quantity, expenditure on goods (2) they do not require an arbitrary assignment of total time available.

Index numbers of the cost of living which incorporate leisure as a commodity can also be defined in terms of the three exogenous variables non-labour income, full income and the after-tax hourly wage rate. These are discussed later. We now turn to some empirical measures of real income and real wages.

4. Index Numbers of Real Income and Real Wages: Canada, 1949-80

Several measures of real income and real wages are presented in this section. The data are annual, and cover the period 1949-80. The base year is 1949. With the exception of the hours worked series the data are from the National Accounts. The commodity breakdown is as follows:

1. Food, beverages and tobacco
2. Clothing and footwear
3. Rent, fuel and power

4. Furniture, furnishings, household equipment and operating costs
5. Medical care and health services
6. Transportation and communications (including net foreign expenditures abroad)
7. Recreation, entertainment, education and cultural services
8. Personal goods and services.

Thus there are eight commodities plus leisure. Net foreign expenditures abroad are included in transportation and communications because these are primarily travel expenditures.

At this level of aggregation, the commodities are themselves quantity indices. Further research with substantially more disaggregate data is planned; however, even at this aggregate level the main consequences of incorporating leisure in measures of economic welfare should emerge.

The hours worked data are from the Labour Force Survey. To be as consistent as possible with the use of national accounts data on income and expenditures, the broadest available measure of hours worked is used. This is actual average hours worked per week in all jobs, the average calculated excluding persons who were not at work during the reference week. The latter ensures that annual hours take account of vacations, etc. Using average hours based on the total employed during the reference week (i.e., including persons who were not at work during the reference week) would overstate annual hours worked unless some alternative correction for vacation time were made. This factor is important because in the period under consideration the increase in leisure time has come about not only from a reduction in hours worked per week but also from a reduction in days worked per year.

Labour income, non-labour income and total expenditure on goods and services (other than leisure) are taken from the income and outlay account of the National Income and Expenditure Accounts. Current transfers to government and other sectors (primarily income taxes) is subtracted from labour income to give after-tax labour income. Dividing this by annual hours worked gives the after-tax wage rate. Non-labour income is net of savings, so that this is non-labour income actually spent. This convention ensures that the budget constraint is met; i.e., that after-tax labour income plus non-labour income sum to total expenditure on goods and services.

Table 1 presents the quantity indices discussed in Section 2 of the paper; i.e., those based on the standard theory of consumer behaviour. These employ only the personal expenditure data from the national accounts; i.e., they ensure that the budget constraint is satisfied by assuming that income equals total personal expenditure.

The four indices Q_L, Q_P, Q_F and Q_T lead to similar conclusions. Real income approximately doubled from 1949 to 1964 and from 1949 to 1976. Since then real income declined slightly.

The table also confirms the expectation that the Min and Max bounds are likely to be too broad to be useful as measures of real income.

Tables 2 and 3 contain index numbers which incorporate income-leisure choice. The various measures based on full income are shown in Table 2. For the full income measures it was assumed that 16 hours per day are available for work or leisure activities.

The index numbers Y_L, Y_P, Y_F and Y_T are very similar in magnitude (indeed, Y_F and Y_T are almost identical throughout the period). They are also consistently lower than their counterparts in Table 1. Real full income rises by, at most, 50 percent from 1949 to 1976, whereas the indices in Table 1 suggest that real income more than quadrupled over this period. The reason for these differences is that the quantity of leisure consumed has not increased at the same rate as other commodities. This is made clear by examining the Min and Max indices in Tables 1 and 2. Adding leisure as a ninth commodity will only change these indices if the ratio of leisure in period t to leisure in the base period is below the minimum or above the maximum for the other eight commodities. As a comparison of the Min columns reveals, leisure is often below the minimum for the other commodities (starred years in Table 2), and when it is not it is close to the minimum. This tends to reduce the growth in the full income indices. For example, Y_T is a weighted average of the ratios of current to base period consumption for each commodity. Adding leisure (consumption of which has been growing at a below average rate) reduces the average.

The economic factors underlying the differences in the growth rates of real income (without allowing for leisure) and real full income are as follows. For the eight purchased commodities, the income effect of rising real wages dominates any substitution effects due

to changes in relative prices. Thus substantial increases in quantities consumed are observed over this period. However, for leisure the income effect of rising real wages is offset by the substitution effect due to the increase in the price of leisure relative to other commodities. Thus the growth in the consumption of leisure time is significantly lower than that of other commodities.

TABLE 1. Index Numbers of Real Income without Accounting for Leisure

Year	Q_L	Q_P	Q_F	Min x_t^i/x_0^i	Max x_t^i/x_0^i	Q_T
1949	100.0	100.0	100.0	100.0	100.0	100.0
1952	110.3	110.4	110.3	101.2	129.6	110.3
1956	130.7	131.2	131.0	115.0	175.4	127.2
1960	143.1	144.2	143.7	120.1	184.7	143.7
1964	151.9	152.0	151.9	127.5	216.9	152.0
1968	162.8	162.3	162.6	128.6	245.7	162.7
1972	182.5	183.4	182.9	108.4	294.6	183.1
1976	202.2	200.4	201.3	128.4	350.0	201.7
1980	194.6	193.3	194.0	126.8	350.0	194.2

TABLE 2. Index Numbers of the Standard of Living Accounting for Leisure

Year	Y_L	Y_P	Y_F	Y_T	Min x_t^i/x_0^i	Max x_t^i/x_0^i
1949	100.0	100.0	100.0	100.0	100.0	100.0
1952	105.8	105.6	105.7	105.7	101.2	129.6
1956	114.8	113.1	113.9	113.9	103.4*	175.4
1960	120.6	117.9	119.2	119.3	104.4*	184.7
1964	124.9	120.5	122.7	122.6	105.6*	216.9
1968	130.7	124.2	127.4	127.3	107.9*	245.7
1972	139.3	129.6	134.2	134.4	108.4	294.6
1976	150.1	134.6	142.1	141.6	112.9*	350.0
1980	147.5	134.5	140.9	140.4	113.9*	350.0

TABLE 3. Index Numbers of Real Wages and Real Non-Labour Income

Year	W_L	W_P	RWS	M_L	M_P
1949	100.0	100.0	100.0	100.0	100.0
1952	115.8	115.8	109.2	285.2	–74.3
1956	140.5	144.0	125.6	568.7	–63.1
1960	157.3	170.1	131.4	752.6	–88.2
1964	170.7	184.2	143.4	890.1	–60.3
1968	190.7	210.9	160.7	1,075.7	–51.4
1972	217.6	265.5	186.4	1,349.0	–52.6
1976	261.7	329.5	214.4	1,689.3	–39.0
1980	256.6	341.5	205.4	1,608.7	–51.0

TABLE 4. Index Numbers of Real Consumption Expenditure

Year	W_L	E_L	Q_L	Q_P
1949	100.0	100.0	100.0	100.0
1952	114.6	114.8	110.3	110.4
1956	137.6	137.2	130.7	131.2
1960	152.8	152.7	143.1	144.2
1964	164.9	163.9	151.9	152.0
1968	183.0	180.6	162.8	162.3
1972	207.5	205.4	182.5	183.4
1976	246.8	239.6	202.2	200.4
1980	241.9	233.0	194.6	193.3

Table 3 contains index numbers of real wages and real non-labour income. The expectation that M_P will often be negative and that the rate of increase of M_L will tend to be large is confirmed.

The W_L and W_P indices indicate that real wages have increased more than would be judged on the basis of traditional "real wage" measures that do not incorporate income-leisure choice. RWS is a comparable traditional real wage measure; it equals real wages and salaries and supplementary labour income per employee divided by the implicit price index for consumer expenditure.

Table 4 contains the Diewert consumption expenditure indices, and the comparable "real income" indices which do not incorporate leisure. As is clear from (69) and (70), the rate of growth of economic welfare is understated by the usual measures which do not account for the decline over time in hours worked. The magnitude of the understatement is indicated in Table 4. The growth in economic welfare was approximately 240 percent over the period 1949-76 rather than approximately 200 percent as the commonly used measures suggest.

5. Conclusions

This paper has accomplished two objectives; (1) to summarize and extend the economic theory of index numbers of the standard of living and the cost of living when allowance is made for joint consumption and labour supply decisions, and (2) to present some measures of real income and real wages which incorporate labour-leisure choice.

Most of the theoretical results in Section 3 on index numbers of economic welfare which incorporate income-leisure choice have analogous results for index numbers of the cost of living. Thus the the cost of living can be measured using full income, non-labour income, or the after-tax wage rate. The main reason for not including a section on the cost-of-living measures which incorporate leisure is that the concept of the cost of living seems limited to goods and services purchased in the marketplace whereas the concept of economic welfare clearly requires going beyond these and taking account of other factors affecting well-being.

While incorporating leisure time in measures of living standards is an important extension of existing measures, there are other extensions that would be worthwhile. One is to incorporate time spent in household production in addition to time allocated to leisure or working for pay. This third activity is particularly important to incorporate if the household is the basic unit of the analysis. A second is to allow non-labour income and wages to be endogenous; i.e., to incorporate savings decisions and decisions about education, training and other human capital investments.

Two *caveats* seem particularly important to mention. First, the theory of welfare measurement used in the paper applies to a single individual whereas the data employed are for the economy as a whole (albeit on a per capita basis). The welfare measures can be taken to apply to a single randomly selected member of the population. Second, the index number theory assumes that the individual is at a maximum utility position each period, or that the individual is working the desired number of hours. Both assumptions are commonly made, but this does not imply that they are unimportant. Measures which make a correction for the possibility that a randomly selected individual is working fewer weeks per year than desired could possibly be based on the aggregate unemployment rate.

Four different welfare measures which incorporate leisure were examined in the paper: full income, non-labour income, wage rate and consumption expenditure measures. The real wage and real consumption expenditure measures appear to be the most useful. Standard ''real wage'' and ''real income'' index numbers understate the increase in living standards; Tables 3 and 4 provide some indication of the degree of understatement.

Footnotes

[1] Throughout the paper I will refer to constant-utility index numbers as real income or cost-of-living index numbers and their observable counterparts as quantity or price indices.

[2] The term full income was introduced by Becker [1965].

References

Abbott, M. and O. Ashenfelter [1976]. "Labour Supply, Commodity Demand and the Allocation of Time", *Review of Economic Studies* 43, October 1976, pp.389-411.

Becker, G.S. [1965]. "A Theory of the Allocation of Time", *Economic Journal* 75, pp.493-517.

Braithwait, S.D. [1980]. "The Substitution Bias of the Laspeyres Price Index: An Analysis Using Estimated Cost-of-Living Indexes", *American Economic Review* 70, March 1980, pp.64-77.

Cleeton, D.L. [1982]. "The Theory of Real Wage Indices", *American Economic Review* 72, March 1982, pp.214-255.

Diewert, W.E. [1976]. "Exact and Superlative Index Numbers", *Journal of Econometrics* Vol. 4, May 1976, pp.115-145.

__________ [1981]. "The Economic Theory of Index Numbers: A Survey", in A. Deaton, ed., *Essays in the Theory and Measurement of Consumer Behavior in Honour of Sir Richard Stone*. Cambridge: Cambridge University Press, 1981.

__________ [1983]. "The Theory of the Cost-of-Living Index and the Measurement of Welfare Change", *Price Level Measurement: Proceedings From a Conference Sponsored by Statistics Canada.*

Lloyd, P.J. [1975]. "Substitution Effects and Biases in Nontrue Price Indices", *American Economic Review* 65, June 1975, pp.301-313.

__________ [1979]. "Constant-Utility Index Numbers of Real Wages: Comment", *American Economic Review* 69, September 1979, pp.682-685.

Noe, N.N. and G. von Furstenberg [1972]. "The Upward Bias in the Consumer Price Index Due to Substitution", *Journal of Political Economy* 80, November-December 1972, pp.1280-1294.

Pencavel, J.H. [1977]. "Constant Utility-Index Numbers of Real Wages", *American Economic Review* 67, March 1977, pp.91-100.

―――――― [1979a]. "Constant-Utility Index Numbers of Real Wages: Revised Estimates", *American Economic Review* 69, March 1979, pp.240-243.

―――――― [1979b]. "Constant-Utility Index Numbers of Real Wages: Reply", *American Economic Review* 69, September 79, pp.686-687.

Samuelson, P.A. and S. Swamy [1974]. "Invariant Economic Index Numbers and Canonical Duality: Survey and Synthesis", *American Economic Review*, 64, pp.566-593.

Usher, D. [1980]. *The Measurement of Economic Growth.* Oxford: Basil Blackwell, 1980.

COMMENTS

John Bossons
Institute for Policy Analysis
University of Toronto

The questions raised by Jack Triplett's paper for this conference apply to the paper by Craig Riddell and I think it is useful to consider Riddell's paper in the light of Triplett's key question: What is the purpose of an index? Before I do so, however, I would like to make one very general point. Much has been said at this conference about the relative advantages of different indexes and the possible benefits associated with different ways of correcting indexes. Particularly when thinking in terms of the applicability of cost of living indexes to tax correction and labour bargaining and in the development of indexed financial contracts, the familiarity of the existing Consumer Price Index is an important advantage. Because changes may reduce the credibility of an accepted index, there may be important costs associated with changing a widely used index. These costs have to be offset against any potential benefits to be derived from such change. I want to stress this point. Looking at it from the viewpoint of using such indexes in tax correction and the development of indexed financial instruments, the stability of the definition of the CPI and its general acceptance are very important.

I have a number of comments on Craig Riddell's paper. A basic point I want to make is that the really interesting question that Riddell has raised is whether there is any gain to be obtained by developing a new ''current period'' cost-of-living index which is more ''complete'' than the Consumer Price Index. The question is a ''second best'' problem, and can be rephrased as follows:

Given that a fully-complete cost-of-living index is ideal but not implemented, is it better to use an incomplete index (a ''subindex'' in the terms used in Diewert's paper to this conference) rather than another? As any student of the second-best problem will appreciate, a ''more complete'' subindex is not necessarily ''better'' in the sense of providing a closer approximation to the unimplemented ideal.

It is important to emphasize first that we are talking about one subindex versus another. No current-period index can be a complete cost-of-living index without taking account of intertemporal allocation. Indeed, contrary to what Riddell has claimed, it is not true that the intertemporal problem arises only if savings is introduced. Even ignoring savings, there are important intertemporal allocation problems inevitably involved in labour-leisure choices. Consequently, I will claim that it is wrong to think that one can easily go from what we can now easily measure, namely a current-period subindex for consumption goods, to anything that is very much further on the road to completeness. I will argue that those difficulties are very relevant in any attempt to define a ''complete'' current-period subindex.

There are several problems that arise in dealing with labor-leisure choices. First, a major complication is introduced by the necessity of dealing with involuntary leisure (i.e. involuntary unemployment). This complication is empirically important. One has to be very careful not to invent implicit welfare measures which ''show'' that welfare was historically maximized in periods such as 1933. An increase in unemployment is an increase in welfare only if one assumes (as in the Lucas-Rapping model) that all unemployment is voluntary. But if one assumes that the temporary Keynesian-type disequilibria can occur in which people get stuck with involuntarily-reduced income and an involuntarily-chosen level of unemployment, a serious problem arises. The introduction of involuntary unemployment necessarily introduces intertemporal planning problems. Involuntary unemployment can be viewed as a constraint on the intertemporal planning problem which reduces welfare. There is a corresponding shadow price which can be viewed as a measure of the value of the availability of work given uncertainty regarding the availability of work. This uncertainty arises both with respect to whether existing job will be continued in subsequent periods or, for somebody who is not now working and is currently consuming leisure, whether leisure taken in the current period can be switched at the consumer's discretion into time spent at work during subsequent periods. The intertemporal labor allocation problem cannot be appropriately characterized without taking risk into account.

Ignoring the risk element introduced by unemployment leaves out a very important component of the problem of constructing a full welfare measure for a current period. This component is correspondingly a component of the problem of defining a ''complete'' current-period cost-of-living index that reflects the consumer's full cost-minimization problem. I want to stress this point in order to emphasize that it is a non-trivial problem to

specify how the currently measured goods price subindex can be expanded into a more complete cost-of-living index. Erwin Diewert has emphasized (in Section 9 of his paper) the difficulties in going beyond a complete current-period subindex to define index numbers that correspond to the intertemporal allocation problem. I'd submit that the difficulties of going from the current CPI index or prices of consumption goods and servies to a ''full'' current-period subindex are equally as great.

Raising the problems created by involuntary unemployment naturally leads to the associated problem of whose leisure time to take as weights in trying to measure a full cost-of-living index. Should a cost-of-living index vary as the age composition of the population changes, changing the average amount of leisure taken through retirement by the population as a whole? Should a cost-of-living index vary with the aggregate level of unemployment, changing as measured labour force participation ratios change in response to the discouraged worker effects? Contrary to the position taken by Riddell, I don't believe that it is straightforward to develop measures of the cost-of-living that incorporate changes in the price of leisure.

Further, I want to argue that incorporating adjustments for changes in labour-leisure choices will for many applications not even be appropriate. Here I want to repeat some points that Jack Triplett made in his paper. For example, in dealing with tax indexation issues, it is necessary to ask what is the purpose of trying to define an index for the purpose of indexing tax rates, exemptions, and so forth. I submit that to a considerable extent what is needed for this purpose is a measure of the trade-off between private and public consumption goods and services. One way of arguing this is to think of this in terms of a utility tree which implies that decisions about the consumption of public goods (including redistributive transfers) are made for given choices by the population about how time expected to be available over their lives will on average be allocated between leisure and time spent in work to create the goods and services which are going to be divided up between private consumption goods and public consumption goods. If one thinks of it that way, then it's not inappropriate to use for such purposes a current-period subindex for consumption goods and services.

I would argue that the errors introduced because of the difficulties that arise in trying to construct a complete single current-period measure decrease the signal-to-noise ratio

for the Index relative to what one now has with the currently-used subindex. Even if it were preferable (ignoring those measurement problems) to use a more complete current-period subindex, for what purposes is Riddell's characterization of the problem relevant?

His key assumptions are that income is endogenous and leisure choice is voluntary. I think Riddell's analysis is very useful for the longer-run questions which he addressed towards the end of his paper, such as the welfare measurement problem earlier introduced by Tobin that arises in correcting the national accounts to obtain a measure that more appropriately measures long-run changes in welfare. However, this analysis of secular trends in welfare is a very different kind of issue from the application that is implied by using an index for purposes of tax corrections, wage escalation, and other short run measurement problems.

Now I would like to make a couple of comments on some broader questions which are partly stimulated by Jack Triplett's paper as well as by Riddell's paper. I want to argue that a very important application of the consumer price index is in its potential role in correcting the tax system for inflation and in the development of indexed financial instruments. Here the problem is to obtain a "correct" measure of real income from capital, which can alternatively be defined as obtaining a measure of the real trade-off between future and present consumption of goods and services.

I think the simplest way to characterize the relevant consumer optimization problem in this application is in a life cycle context in which the most important price is the price index for consumption goods during retirement, at which point in the life cycle a person can be assumed to be 100 percent engaged in leisure. With this simplification, the relevant consumer intertemporal allocation problem can be characterized as one in which the leisure/labor choice is not relevant. While obviously excluding an important aspect of the complete intertemporal problem, I believe this characterization is relevant for the purpose of getting at the critical indexing problem that applies when one is defining a "correct" definition of taxable income from capital, the appropriate definition of income from capital when one is constructing an indexed financial instrument, of the appropriate index to be used in indexing a life annuity or pension. I would consequently argue that a current-period subindex for consumption goods is very close to the correct index that should be used for

this purpose, so that the consumer price index or some variant is, in fact, ''the right index'' for this purpose.

The example provided by the application to defining a ''correct'' measure of income from capital for purposes of taxation or for defining financial contracts and pensions in real terms illustrates the importance of the points emphasized by Triplett. The appropriate characterization of the relevant consumer allocation problem differs depending on the application, and this has important implications for the choice of index. Depending on the application, a current-period subindex for consumption goods and services may be a better measure than a ''complete'' cost-of-living index even ignoring the measurement problems that would arise if the ''complete'' index were to be used. Where this is the case, the ''second best'' problem analyzed in Riddell's paper becomes irrelevant.

PREFERENCE DIVERSITY AND AGGREGATE ECONOMIC COST-OF-LIVING INDEXES[1]

Charles Blackorby and David Donaldson
Department of Economics
University of British Columbia

SUMMARY

For a single household, an economic cost-of-living index compares the minimum expenditure needed to achieve a particular indifference curve under two different price regimes.

We are interested in a price index for a group of households rather than a single one. Further, it is likely that preferences, even if homothetic, are different for different households. It would be useful to be able to combine the individual price indexes into an aggregate price index that is itself capable of being given a reasonable economic interpretation. Two general possibilities exist, the non-optimizing and optimizing approaches.

The non-optimizing approach attempts to find conditions under which aggregate demand behaviour could be generated by a "representative" household. Preferences of this representative household are then used to construct an expenditure function, which is, in turn, used to construct a cost-of-living index. The alternative optimizing procedure is to construct a social expenditure function based on a (Bergson-Samuelson) social-welfare function and individual preferences. This social expenditure function is then used to construct a social cost-of-living index. In this paper, we pursue both of these procedures, and attempt to find conditions under which the resulting aggregate cost-of-living index is reference-level free.

RÉSUMÉ

Pour un ménage unique, un indice économique du coût de la vie compare la dépense

minimale nécessaire pour obtenir une courbe d'indifférence particulière avec deux régimes de prix différents.

Nous nous intéressons à un indice de prix pour un groupe de ménages plutôt qu'un seul ménage. De plus, il est probable que les préférences, même si elles sont homothétiques, sont différentes pour différents ménages. Il serait utile de pouvoir combiner les indices individuels de prix en un indice agrégé de prix qui serait lui-même capable de fournir une interprétation économique raisonnable. Il existe deux possibilités générales, la technique de non-optimisation et la technique d'optimisation.

La technique de non-optimisation tente de trouver des conditions qui permettraient de générer le comportement de demande globale à partir d'un ménage « représentatif ». On utilise alors les préférences de ce ménage représentatif pour construire une fonction de dépenses, laquelle, à son tour, est utilisée pour établir un indice du coût de la vie. La procédure d'optimisation, pour sa part, consiste à construire une fonction des dépenses sociales basées sur une fonction de bien-être social (Bergson-Samuelson) et les préférences individuelles. On utilise alors cette fonction de dépenses sociales pour mettre au point un indice social du coût de la vie. Dans ce texte, nous utilisons ces deux techniques et essayons de fixer les conditions auxquelles l'indice agrégé du coût de la vie qui en résulte est indépendant du niveau d'utilité de référence.

Dans la section 2, nous discutons l'existence d'un consommateur représentatif et d'un indice correspondant. D'abord nous examinons le cas des prix obtenus de sources exogènes et des distributions arbitraires du pouvoir d'achat. L'indice résultant est indépendant du niveau d'utilité de référence du ménage représentatif si, et seulement si, l'économie est composé de ménages qui ont des préférences identiques homothétiques. Nous analysons alors la même question dans un domaine de prix qui est potentiellement moins exigeant. Nous permettons d'abord la détermination endogène, dans une situation d'équilibre général, de quelques-uns des prix et ensuite de tous. Malheureusement, ce relâchement ne produit pas de solutions nouvelles, car il nous laisse encore avec des préférences identiques homothétiques. Nous examinons également rapidement une situation dans laquelle les fonctions de dépenses du ménage dépendent d'attributs autres que les niveaux d'utilité. Encore une fois, ce relâchement ne conduit à aucune nouvelle généralisation; tous les ménages

doivent avoir des préférences identiques homothétiques si l'on veut que l'indice du coût de la vie soit indépendant du niveau de référence.

Dans la section 3, nous nous tournons vers l'approche de maximisation. Conformément à Pollak, nous construisons une fonction de dépenses sociales. C'est la dépense globale minimale, qui permet la redistribution du pouvoir d'achat en montant global, et qui est nécessaire pour atteindre un niveau donné de bien-être social. Nous l'utilisons alors pour construire un indice « social » du coût de la vie, avec l'exigence que l'indice soit indépendant du niveau de référence (indépendant du niveau de bien-être social choisi). Nous notons que des préférences homothétiques identiques produisent un indice social du coût de la vie indépendant du niveau de référence, mais nous nous préoccupons principalement de la possibilité que différents ménages (ou différents sous-groupes de la population) puissent avoir des préférences différentes. Nous analysons des fonctions linéaires arbitraires de bien-être social (concentrant ainsi notre attention sur des fonctions de bien-être social individuellement additives). Notre résultat principal est qu'il est possible de mettre au point un indice du coût de la vie social qui soit indépendant du niveau de référence lorsque les ménages ont des préférences homothétiques mais non identiques. Cet indice global dépend, en général, des poids de la fonction de bien-être social. On peut facilement calculer ces indices à partir des fonctions de dépenses unitaires distinctes, ($\{\Pi^h(p)\}$) et nous présentons des formules. Nous discutons la relation entre nos indices et la moyenne pondérée également et par partie des indices individuels de prix.

Il est également possible de se demander si on peut choisir des fonctions de bien-être social permettant de rationaliser la distribution existante du pouvoir d'achat, et nous montrons comment cela peut être réalisé. Nous discutons de plus, en conclusion, de la possibilité de choisir des poids pour trouver des poids sociaux implicites dans le choix d'un indice particulier du coût de la vie.

1. Introduction

For a single household, an economic cost-of-living index compares the minimum expenditure needed to achieve a particular indifference curve under two different price regimes. More specifically, let the preferences of a particular household (h) be represented by the

expenditure function e^h[2], so that $e^h(u_h, p)$ is the minimum expenditure needed by the household to achieve the level of welfare indexed by u_h at prices $p = (p_1, \ldots, p_m)$. A (Konüs-type) cost-of-living index for household h is defined to be

$$I^h(p^1, p^0, \bar{u}_h) = \frac{e^h(\bar{u}_h, p^1)}{e^h(\bar{u}_h, p^0)}, \tag{1}$$

the ratio of minimum expenditures at "comparison" prices p^1 to minimum expenditures at "reference" prices p^0. This index is homogeneous of degree one in comparison prices and homogeneous of degree minus one in reference prices. The reference level of well-being is indexed by $\bar{u}_h$. It may refer to a given indifference surface of the household, or it may be the level of utility generated by a particular commodity bundle $\bar{x}^h$, with

$$\bar{u}_h = U^h(\bar{x}^h) \tag{2}$$

where U^h is a (direct) utility function corresponding to e^h.[3]

For many purposes, however, we want the index I^h to be independent of the reference level of utility, $\bar{u}_h$. For this to be true, it is necessary and sufficient that the preferences of household h be homothetic. That is, U^h can be written as

$$U^h(x^h) = \overset{*}{U}^h(\bar{U}^h(x^h)), \tag{3}$$

where $\overset{*}{U}^h$ is increasing and $\bar{U}^h$ is positively linearly homogeneous. Equivalently, the preferences are homothetic if and only if the expenditure function e^h can be written[4] as

$$e^h(u_h, p) = \Pi^h(p)\,\phi^h(u_h) \tag{4}$$

where Π^h is positively linearly homogeneous and ϕ^h is increasing. In this case the index I^h is reference-level free,[5] and

$$I^h(p^1, p^0, \bar{u}_h) = \frac{\Pi^h(p^1)}{\Pi^h(p^0)}. \tag{5}$$

In general, we are interested in a price index for a group of households rather than a single one. Further, it is likely that preferences, even if homothetic, are different for different households. It would be useful to be able to combine the individual price indexes into an aggregate price index that is itself capable of being given a reasonable economic interpretation. Two general possibilities exist, the non-optimizing and optimizing approaches.

The non-optimizing approach attempts to find conditions under which aggregate demand behaviour could be generated by a "representative" household (Gorman [1953, 1961]). Preferences of this representative household are then used to construct an expenditure function, which is, in turn, used to construct a cost-of-living index. The alternative optimizing procedure (Pollak [1980, 1981]) is to construct a social expenditure function based on a (Bergson-Samuelson) social-welfare function and individual preferences. This social expenditure function is then used to construct a social cost-of-living index.

In this paper, we pursue both of these procedures, and attempt to find conditions under which the resulting aggregate cost-of-living index is reference-level free. In Section 2, we discuss the existence of a representative consumer and the corresponding index. First we look at the case of exogenously given prices and arbitrary distributions of purchasing power. The resulting index is independent of the reference level of utility of the representative household if and only if the economy is composed of households which have identical homothetic preferences. We then investigate the same question with a potentially less demanding price domain. We allow first some, and then all prices to be determined endogenously in a general equilibrium setting. Unfortunately, this relaxation produces no new solutions, leaving us again with identical homothetic preferences. We also investigate briefly (following Lau [1977], Jorgenson, Lau and Stoker [1981, 1982]) a situation in which household expenditure functions depend on attributes other than utility levels. Again, this relaxation leads to no new generalizations; all households must have identical homothetic preferences if the aggregate cost-of-living index is to be reference-level free.

In Section 3, we turn to the maximizing approach. Following Pollak [1980, 1981], we construct a social expenditure function. It is the minimum aggregate expenditure, allowing lump-sum redistributions of purchasing power, necessary to achieve a given level of social welfare. We then use it to construct a **social** cost-of-living index and ask that the index be reference-level free (independent of the chosen level of social welfare). We note that identical homothetic preferences produce a reference-level-free social cost-of-living index, but we are principally concerned with the possibility that different households (or population subgroups) may have different preferences. We consider arbitrary linear social-welfare functions (thus restricting our attention to additively separable social-welfare functions). Our main result is that it is possible to find a social cost-of-living index that is reference-level free when households have homothetic but non-identical preferences. This aggregate index depends, in general, on the weights in the social-welfare function. These indexes are easily computed from individual unit expenditure functions, ($\left\{\Pi^h(p)\right\}$) and we provide formulas. We discuss the relationship of our indexes to the Prais [1959] – Muellbauer [1974] – Nicholson [1975] – Pollak [1980] share-weighted average and equally-weighted average of individual price indexes.

It is also possible to ask if the social-welfare function could be chosen to rationalize the existing distribution of purchasing power, and we show how it can be done. We further discuss, in the concluding remarks, the possibility of choosing weights to find the social weights implicit in the choice of a particular cost-of-living index.

2. Representative Households

In this section, we explore the problem of the existence of a representative household under different price regimes. In the first case, prices are exogenously given, in the second, some or all of the prices are determined endogenously, and in the third, households are identified by special characteristics such as age of head, region and so on.

2.1 Exogenous Prices

Here, we review the well-known case explored by Gorman [1953]. We present it in some detail to make further exposition clear.

Each household faces prices p with income $y_h > 0$, and we may write aggregate demand for good j as

$$x_j(p, y_1,\ldots,y_H) = \sum_h x_j^h(p, y_h), \qquad (6)$$

$j = 1,\ldots,m$. This places no restrictions on the aggregate demand functions. However, we are interested in finding conditions under which the aggregate demands of (6) could be generated by a single rational agent whose income is $Y = \sum_h y_h$. Thus, we require

$$x_j(p, y_1,\ldots,y_H) = \bar{x}_j(p, Y) = \sum_h x_j^h(p, y_h), \qquad (7)$$

for a demand function $\bar{x}_j$, $j = 1,\ldots,m$. Since the representative household is assumed to be rational, it must have an expenditure function. If its indifference curves are indexed by w, then we must have, since $Y = \sum_h y_h$,

$$e(w, p) = \sum_h e^h(u_h, p) \qquad (8)$$

for all p. By setting $p = (1,\ldots,1) = 1_m$ we get

$$e(w, 1_m) = \sum_h e^h(u_h, 1_m), \qquad (9)$$

and with a simple normalization of w,

$$w = \sum_h \psi^h(u_h). \qquad (10)$$

Since equation (8) is an identity, it may be differentiated, using Shephard's lemma, to obtain

$$\hat{x}_j(w, p) = \sum_h \hat{x}_j^h(u_h, p) \qquad (11)$$

for all p, where $\hat{x}_j$ and $\{\hat{x}_j^h\}$ are the compensated demand functions for good j for the representative and given households. Substitution of indirect utility functions[6] into (11) yields (7), thus reducing the problem to finding the conditions under which

$$e\left(\sum_h \psi^h(u_h), p\right) = \sum_h e^h(u_h, p) \tag{12}$$

can hold for all p and for all $u = (u_1, \ldots, u_H)$.

We want to know what restriction (12) places on the individual and representative expenditure functions. The solution to this problem is well known and due originally to Gorman [1953]. There exists a representative household if and only if each household's expenditure function may be written as

$$e^h(u_h, p) = \Pi(p)\,\psi^h(u_h) + \beta^h(p), \tag{13}$$

and

$$e(w, p) = \Pi(p)\,w + \beta(p), \tag{14}$$

where w is given by (10) and

$$\beta(p) := \sum_h \beta^h(p). \tag{15}$$

This means, since Π is independent of h in (13), that the Engel curves for the households must be parallel straight lines. The aggregate cost-of-living index

$$I(p^1, p^0, \bar{w}) = \frac{e(\bar{w}, p^1)}{e(\bar{w}, p^0)} = \frac{\Pi(p^1)\,\bar{w} + \beta(p^1)}{\Pi(p^0)\,\bar{w} + \beta(p^0)}, \tag{16}$$

and, in order for it to be reference-level free, we must have

$$\beta(p) = \sum_h \beta^h(p) = c\,\Pi(p), \tag{17}$$

for some real number c. In this case,

$$I(p^1, p^0, \bar{w}) = \frac{\Pi(p^1)}{\Pi(p^0)}. \tag{18}$$

The representative household's preferences are homothetic, but (strictly speaking) each household does not necessarily have homothetic preferences. (14) allows "committed" expenditures by households, and (17) requires that non-proportional committed expenditures sum to zero. Thus some committed expenditures must be positive and others **negative** (not allowing (14) to hold everywhere) if (17) holds without all committed expenditures zero. For this reason, and the fact that these differences in taste are very sensitive to changes in the number of households, we choose to ignore the differences in taste associated with these expenditures and set them all equal to zero. In this case each household has identical homothetic preferences, and its reference-level-free cost-of-living index is the appropriate one for the group.

This result is very restrictive. It results from the requirement that the distribution of income (and hence of utilities) be arbitrary, at least in some open region of income space, and the requirement that all prices be exogenous.

2.2 Endogenously Determined Prices

We therefore attempt to generalize Gorman's result by allowing some or all of the prices to be determined endogenously by the interaction of supply and demand. To do this, we partition the list of commodities $J = (1,\ldots,m)$ into two groups; $\hat{J}$ is the set of commodities whose prices are exogenously given, and $\overset{o}{J}$ is the set whose prices are endogenously determined. Thus

$$J = (\hat{J}, \overset{o}{J}) \tag{19}$$

and we write the corresponding prices as

$$p = (\hat{p}, q). \tag{20}$$

Since, by Shephard's lemma, the price derivatives of expenditure functions are compensated demand functions, we may write our requirement as

$$e_j (w, \hat{p}, q) = \sum_h e_j^h (u_h, \hat{p}, q), \ \ j \in \hat{J} \tag{21}$$

and

$$e_j (w, \hat{p}, q) = \sum_h e_j^h (u_h, \hat{p}, q) = S^j (\hat{p}, q), \ j \in \overset{o}{J}, \tag{22}$$

where e_j and e_j^h are the derivatives of e and e^h with respect to the j^{th} price. S^j is the supply function for the j^{th} good. (21) and (22) must hold only for price-utility combinations that are compatible with the price determination in (22). We assume that (22) can be solved to get

$$q_j = Q^j (\hat{p}, w) , \ j \in \overset{o}{J}, \tag{23}$$

and we assume that the solution is unique (in the case when $\hat{J}$ is empty, we select a *numeraire*). We could, alternatively, solve this system of equations as

$$w = \Gamma^j (\hat{p}, q_j) \ , \ \ j \in \overset{o}{J}. \tag{24}$$

Setting $\hat{p} = (1,\ldots,1)$ in (23) allows $\hat{p}$ and q to be substituted out of (21) and (22). We assume that these equations can be solved uniquely for w as

$$w = W(u) = W(u_1,\ldots, u_H). \tag{25}$$

W must be increasing in its arguments in order that aggregate expenditure respond positively to individual expenditures.

Equations (21) and (22) imply, using Euler's theorem and the positive linear homogeneity of expenditure functions in prices, that

$$e\,(w, \hat{p}, q) = \sum_h e^h\,(u_h, \hat{p}, q), \tag{26}$$

or that (using (23) and (25))

$$e(W(u), \hat{p}, Q\,(\hat{p}, W\,(u))) = \sum_h e^h(u_h, \hat{p}, Q(\hat{p}, W(u))). \tag{27}$$

for all $\hat{p}$ and u, where

$$Q(\hat{p}, w) = \left(\{Q^j(\hat{p}, w)\}_{j \in \overset{o}{J}}\right). \tag{28}$$

Equation (27) is similar to equation (12), but there are several differences. First, we cannot be sure that W is additive in household utilities. Second, the domain of (26) is restricted by the endogenous pricing. In fact, (21) and (22) are **not** implied by (26) unless there are no endogenous prices. We require (21) and (22) to be satisfied, and, therefore, satisfaction of (26) is necessary but not sufficient. It would seem intuitively clear, however, that these new conditions would allow more scope for individual variability than before. That this is **not** the case is stated in

Theorem 1: Given our regularity conditions, (21) and (22) hold for all $\hat{p}$, u if and only if the expenditure functions can be written as

$$e^h(u_h, p) = \Pi\,(p)\,\psi^h\,(u_h) + \beta^h\,(p), \tag{29}$$

and

$$e(w, p) = \Pi(p)\,w + \beta(p), \tag{30}$$

where

$$\beta(p) := \sum_h \beta^h(p) \tag{31}$$

and

$$w = W(u) = \sum_h \psi^h(u_h). \tag{32}$$

Proof: See the appendix.

In the proof, we show that W must satisfy a particular functional restriction, namely that (with a harmless normalization),

$$w = \sum_h F^h(u_h, w), \tag{33}$$

where F^h is increasing in u_h. Subsequently, we show that (21) and (22) require that F^h not depend on w. Thus, no change in preferences from the solution to Gorman's problem has been found. Hence, there exists a reference-level-free aggregate cost-of-living index if and only if all households have identical homothetic preferences. The cost-of-living index is given by (18).

2.3 Households with Multiple Characteristics

In recent years, several articles[7] have investigated aggregate demand equations which are identified by various attributes or characteristics. This allows a representative household to exist with preferences depending on some aggregate values of those characteristics. This has led to representative preferences that are more general than the Gorman restriction on preferences (above) and are easy to implement.

We pursue this notion briefly. Suppose that each household has a level of a characteristic indexed by v_h. Then we require

$$e(w_1, w_2, p) = \sum_h e^h (u_h, v_h, p) \tag{34}$$

for all p, $\{u_h\}$ and $\{v_h\}$. Setting $p = 1_m$ in (34) requires that the aggregate characteristics w_1 and w_2 depend on u and v. We require that

$$w_1 = W^1 (u,v) \tag{35}$$

and

$$w_2 = W^2 (u, v). \tag{36}$$

The solution to this problem is known (Gorman [1978]) to be given by

$$e^h(u_h, v_h, p) = {}^1\Pi (p)\psi^{h1} (u_h, v_h) + {}^2\Pi(p)\psi^{h2} (u_h, v_h) + \beta^h (p), \tag{37}$$

and

$$e(w_1, w_2, p) = {}^1\Pi (p) w_1 + {}^2\Pi (p) w_2 + \beta (p), \tag{38}$$

where

$$w_1 = W^1(u,v) = \sum_h \psi^{h1} (u_h, v_h), \tag{39}$$

$$w_2 = W^2(u,v) = \sum_h \psi^{h2} (u_h, v_h), \tag{40}$$

and

$$\beta(p) := \sum_h \beta^h(p). \tag{41}$$

The aggregate cost-of-living index in this case is

$$\bar{I}\,(p^1, p^0, w_1, w_2) = \frac{{}^1\Pi\,(p^1)\,w_1 + {}^2\Pi\,(p^1)\,w_2 + \beta(p^1)}{{}^1\Pi\,(p^0)\,w_1 + {}^2\Pi\,(p^0)\,w_2 + \beta(p^0)}\,, \tag{42}$$

and we want it to be independent of w_1 and w_2. The conditions for this are

$${}^1\Pi(p) = c_1\,\Pi(p), \tag{43}$$

$c_1 > 0$,

$${}^2\Pi(p) = c_2\,\Pi(p), \tag{44}$$

$c_2 \geq 0$,

$$\beta(p) = c_3\,\Pi(p), \tag{45}$$

for all p. Thus

$$e(w_1, w_2, p) = \Pi(p)\,(c_1\,w_1 + c_2\,w_2 + c_3). \tag{46}$$

In each of these cases, households must have identical homothetic preferences. In constructing aggregate cost-of-living indexes, nothing has been gained.

3. The Social Cost-of-Living Index

Pollak [1980, 1981] has proposed a procedure for finding an aggregate cost-of-living index based on a social-welfare function. It is derived from a social expenditure function – the minimum expenditure required (with lump-sum redistributions) to achieve a given level of social welfare.

Following the usual practice, we assume that social welfare is

$$w = W(u) \tag{47}$$

and that W is continuous, differentiable, and quasi-concave. We now need the assumption that each expenditure function is **strictly convex** in u_h. This rules out degenerate solutions, and is equivalent to the assumption that each household's direct utility function has a strictly concave representation.[8]

We define the social expenditure function as

$$\bar{e}\,(w, p) = \inf_u \left\{ \sum_h e^h(u_h, p) \;\middle|\; W(u) \geq w \right\}. \tag{48}$$

We want to limit our attention to additively separable social-welfare functions, where

$$w = \sum_h a_h \, \psi^h \, (u_h), \tag{49}$$

and each ψ^h is concave. A simple normalization[9] of individual utilities makes this into

$$w = \sum_h a_h \, u_h. \tag{50}$$

We assume that no a_h is negative and that at least one is positive. In this case, the social expenditure function is

$$e(w, p, a) = \inf_u \left\{ \sum_h e^h \, (u_h, p) \;\middle|\; \sum_h a_h \, u_h \geq w \right\}, \tag{51}$$

and it depends on the level of social welfare, w, the price vector p, and the vector of social weights a.

The social cost-of-living index is defined in the usual manner, using the social expenditure function e. The ratio of expenditures necessary to achieve social-welfare level w in the two price regimes is

$$I(p^1, p^0, w, a) = \frac{e(w, p^1, a)}{e(w, p^0, a)}. \tag{52}$$

We would like to discover the conditions under which this cost-of-living index is independent of w. It is easy to show that, for this to be true it is necessary and sufficient that the social expenditure function can be written as

$$e(w, p, a) = \Lambda\,(p,a)\; T\,(w, a), \tag{53}$$

implying that

$$I(p^1, p^0, w, a) = \frac{\Lambda(p^1, a)}{\Lambda(p^0, a)}. \tag{54}$$

The social expenditure function must be decomposable into two functions. One depends on prices and social weights a, the other depends on w and the social weights. In general, the social cost-of-living index, even if reference-level free, depends on a, and in the special case that it does not, individual preferences will again prove to be identical and homothetic.

The domain of u_h in e^h is an interval of the real line (or, perhaps, all of it), and we denote this as D^h. Again we rule out precommitted expenditures (as in Section 2) and require

$$\inf\left\{e^h(u_h, p) \mid u_h \in D^h\right\} = 0\,.^{10} \tag{55}$$

By setting $a_g = 0$ for all $g \neq h$, $a_h = 1$, we note that (55) requires

$$e(w,p,(0,\ldots,1,\ldots,0)) = e^h\,(w, p). \tag{56}$$

Thus, satisfaction of (53) for all a requires

$$e^h(u_h,p) = \Pi^h\,(p)\,\psi^h\,(u_h). \tag{57}$$

Therefore, a necessary condition (given (55)) for reference-level freedom of the social cost-of-living index is that households have (possibly non-identical) homothetic preferences. Further, if a single household is the only one that counts in social welfare, its cost-of-living index is the social cost-of-living index.

To proceed with the analysis of possibilities for individual preferences, we prove a simple social duality result. We denote the optimal value of u_h in (51) as u_h^*, so that, whenever these values exist,

$$
\begin{aligned}
e(w,p,a) &= \min\left\{\sum_h e^h(u_h, p) \;\middle|\; \sum_h a_h u_h \geq w\right\} \\
&= \sum_h e^h(u_h^*, p).
\end{aligned}
\tag{58}
$$

Lemma 1: If $u_h^* \in \operatorname{int} D^h$ for all h, then

$$
u_h^* = -\frac{e_{a_h}(w,p,a)}{e_w(w,p,a)} \tag{59}
$$

where the subscripts denote partial derivatives.

Proof: See the appendix.

We may use the special structure on e(w,p,a) in (53) to place structure on the optimal level of utilities.

$$
\begin{aligned}
u_h^* &= -\frac{\Lambda_{a_h}(p,a)\,T(w,a) + \Lambda(p,a)\,T_{a_h}(w,a)}{\Lambda(p,a)\,T_w(w,a)} \\
&= -\left[\frac{\Lambda_{a_h}(p,a)}{\Lambda(p,a)}\right]\left[\frac{T(w,a)}{T_w(w,a)}\right] - \left[\frac{T_{a_h}(w,a)}{T_w(w,a)}\right] \\
&:= \alpha^h(p,a)\,\Gamma(w,a) + \Phi^h(w,a)
\end{aligned}
\tag{60}
$$

Our main results are derived from (60) and the first-order conditions for problem (51).

Several different possibilities arise, depending on the nature of the functions in (53) and (60). If the function is independent of a, the cost-of-living index is independent of the social weights. This suggests immediately that individual preferences must be identical as well as homothetic. Thus, we are taken back to the representative household of Section 2.

There are other solutions, however. They allow different homothetic preferences for the households. They correspond to the cases where $\Gamma(w, a)$ is independent of w, and where $\Phi^h(w, a)$ is independent of w. Our main results are contained in

Theorem 2: Given our regularity conditions,[11] the social cost-of-living index is reference-level free for all $a > 0$, if and only if individual households' expenditure functions can be written as one of

Case 1:

$$e^h(u_h, p) = \Pi(p)\psi^h(u_h), \tag{61}$$

where ψ^h is a strictly convex function, $h = 1,\ldots H$,

Case 2:

$$e^h(u_h, p) = \Pi^h(p) \exp\{u_h\} \tag{62}$$

where D^h is $\mathbb{R}$, $h = 1,\ldots,H$,

or

Case 3:

$$e^h(u_h, p) = \Pi^h(p)\left[r(u_h - c_h)\right]^{\frac{1}{r}}, \tag{63}$$

where $D^h = \{u_h \mid u_h \geq c_h\}$, $0 < r < 1$,

and $D^h = \{u_h \mid u_h \leq c_h\}$, $r < 0$, $h = 1,\ldots,H$.

Corresponding to these household preferences, the social cost-of-living indexes are

Case 1:

$$I(p^1, p^0, w, a) = \frac{\Pi(p^1)}{\Pi(p^0)}, \tag{64}$$

Case 2:

$$I(p^1, p^0, w, a) = \underset{h}{X}\left[\frac{\Pi^h(p^1)}{\Pi^h(p^0)}\right]^{\bar{a}_h}, \tag{65}$$

where $\bar{a}_h := a_h / \sum_g a_g$, and

Case 3:

$$I(p^1, p^0, w, a) = \frac{\left[\sum_h a_h^{\frac{1}{1-r}} \Pi^h(p^1)^{\frac{-r}{1-r}}\right]^{\frac{1-r}{-r}}}{\left[\sum_h a_h^{\frac{1}{1-r}} \Pi^h(p^0)^{\frac{-r}{1-r}}\right]^{\frac{1-r}{-r}}} \tag{66}$$

Proof: See the appendix.

These results merit some discussion. Case 1 is identical homothetic preferences, and the social cost-of-living index is the common individual index. Of course, the social cost-of-living index is the common individual one for any social-welfare function whatever. Cases 2 and 3 allow non-identical preferences. In Case 2, the social cost-of-living index is a Cobb-Douglas function of the individual price indexes. If everyone receives an equal weight, then the social cost-of-living index is the geometric mean of the individual indexes. This contrasts with the "democratic" price index of Prais and Nicholson (generalized by Muellbauer and Pollak). It is an arithmetic mean of individual price indexes. The index in Case 3 is a ratio of weighted means of order (-r/(1-r)) of the individual "unit" cost functions $\Pi^h(p)$. (-r/(1-r)) can take on all real values between $-\infty$ and one excluding zero). The limiting case of zero is just Case 2. As r approaches $-\infty$, the index becomes

$$\frac{\sum_h \Pi^h(p^1)}{\sum_h \Pi^h(p^0)} = \frac{\frac{1}{H}\sum_h \Pi^h(p^1)}{\frac{1}{H}\sum_h \Pi^h(p^0)}, \tag{67}$$

the ratio of means of unit costs. This is closely related to, but not equal to the democratic price index.

The above indexes may be rewritten in terms of optimal expenditure shares, $\{s_h^*\}$. For household h, in Case 2,

$$s_h^* = \frac{a_h}{\sum_g a_g} = \bar{a}_h \ . \tag{68}$$

Thus, the social cost-of-living index can be rewritten as

$$I(p^1, p^0, w, a) = \underset{h}{X} \left(\frac{\Pi^h(p^1)}{\Pi^h(p^0)} \right)^{s_h^*} \tag{69}$$

In Case 3, the optimal expenditure share for household h depends on the price vector p as well as the set of weights, and

$$s_h^* = S^h(p, a) = \frac{a_h^{\frac{1}{1-r}} \; \Pi^h(p)^{-\frac{r}{1-r}}}{\sum_g a_g^{\frac{1}{1-r}} \; \Pi^g(p)^{-\frac{r}{1-r}}}, \tag{70}$$

and the social cost-of-living index can be shown to be

$$I(p^1, p^0, w, a) = \frac{\left[\sum_h S^h(p^1, a)^{(1-r)} \; \Pi^h(p^1)^r\right]^{\frac{1}{r}}}{\left[\sum_h S^h(p^0, a)^{(1-r)} \; \Pi^h(p^0)^r\right]^{\frac{1}{r}}}. \tag{71}$$

(70) and (71) suggest that the indexes we have proposed are related to the "plutocratic" index of Prais and Nicholson (the base-period-share-weighted average of individual Laspeyres indexes) and its generalizations by Pollak (the same average of arbitrary household indexes).[12] However, the mean that appears here is not an ordinary arithmetic mean, and the unit cost functions rather than the individual indexes are used in the averages.

Practical use of these indexes demands that an important question be discussed. That is the question of choosing a and r (Case 2 corresponds to r = 0). r is an "inequality-aversion" parameter. This may be most easily seen by computing the Bergson-Samuelson indirect utility functions for Cases 2 and 3 (we set $c_h = 0$ in Case 3 for simplicity). In each case, we simply find $\Sigma a_h u_h$ with the indirect utility functions substituted in. Since in Case 2,

$$e^h(u_h, p) = \Pi^h(p) \exp\{u_h\} = y_h \longleftrightarrow u_h = \log \frac{y_h}{\Pi^h(p)}, \tag{72}$$

we have found the individual indirects, and $\Sigma\, a_h u_h$ is equivalent to

$$B(p, y_1,\ldots,y_H) = \sum_h a_h \log\left(\frac{y_h}{\Pi^h (p)}\right) \tag{73}$$

$$= \log\left[\underset{h}{X}\left(\frac{y_h}{\Pi^h (p)}\right)^{a_h}\right].$$

In Case 3, a similar computation yields

$$B(p, y_1,\ldots,y_H) = \sum_h \frac{a_h}{r}\left(\frac{y_h}{\Pi^h(p)}\right)^r, \; r < 1, r \neq 0. \tag{74}$$

Thus, the Bergson-Samuelson indirects are the quasi-concave CES-Cobb- Douglas family of functions, the (weighted) means of order r ($r < 1$), and the independent variables are individual real incomes. $r = 1$ represents no inequality aversion, $r = -\infty$ is "maximin" – complete inequality aversion. The weights a_h can be set normatively, and in this case they would presumably be equal (differences in need ought to appear in the unit cost functions $\{\Pi^h (p)\}$. In this case, the social cost-of-living indexes can be computed directly from (65) and (66), presumably for several values of r.

On the other hand, it is possible (for any r) to choose the a's to rationalize any pattern of expenditure shares. Thus it would be possible to use actual expenditure shares to justify a choice of a's. This presumes implicitly that the government is maximizing a social-welfare function with (possibly peculiar) weights. If this approach is taken, the social cost-of-living indexes can be computed directly from (69) and (71).

This procedure suggests another exercise. Suppose that we have estimated a unit cost function for each income class in the economy, treating each group as if it consisted of identical households. We can then try different values of a and r to see which ones best fit the actual cost-of-living indexes constructed by government agencies. In this way, another "inverse-optimal" problem could be solved.

4. Conclusion

We have investigated two general procedures for finding reference-level-free aggregate cost-of-living indexes. The non-maximizing approach involves the search for a representative consumer, and we have shown that this approach results in an unhelpful solution. All households must have identical homothetic preferences, with the aggregate index equal to the individual indexes.

The maximizing framework allows more variability in individual preferences since the distribution of purchasing power is adjusted to suit the dictates of a social-welfare function in finding the social cost-of-living index. We have investigated the case of additively separable social-welfare functions only, and have found all the sets of individual preferences that will produce reference-level-free social cost-of-living indexes for arbitrary social weights. Individual households must have homothetic preferences, but they may differ from household to household. The social cost-of-living indexes are ratios of weighted means of individual unit cost functions. These weighted means are all in the CES-Cobb-Douglas family and contain, in addition to the social weights, an inequality-aversion parameter.

Appendix

Theorem 1: Given our regularity conditions, (21) and (22) hold for all $\hat{p}$, u if and only if the expenditure functions can be written as

$$e^h(u_h, p) = \Pi(p)\, \psi^h(u_h) + \beta^h(p), \tag{29}$$

and

$$e(w, p) = \Pi(p)\, w + \beta(p), \tag{30}$$

where

$$\beta(p) := \sum_h \beta^h(p) \tag{31}$$

and

$$w = W(u) = \sum_h \psi^h (u_h). \tag{32}$$

Proof: We consider two cases, first the case where $|\overset{o}{J}| = 1$, and second the case where $|\overset{o}{J}| = m$.

Case 1: Equation (27) may be differentiated with respect to u_h to yield

$$e_w W_h + e_q \frac{\partial q}{\partial u_h} = e_u^h + \sum_g e_q^g \frac{\partial q}{\partial u_h}. \tag{a.1}$$

Using (22), this becomes

$$e_w(W(u), \hat{p}, Q(W(u), \hat{p}))\, W_h(u) = e_u^h (u_h, \hat{p}, Q(W(u), \hat{p})). \tag{a.2}$$

Since the right-hand side of this equation depends on $\hat{p}$, u_h and w only, so does the left. Consequently, we can write

$$W_h(u) = f^h (u_h, W(u)), \tag{a.3}$$

and

$$e_u^h (u_h, \hat{p}, Q(w, \hat{p})) = e_w(w, \hat{p}, Q(w, \hat{p}))\, f^h(u_h, w). \tag{a.4}$$

Substituting $q = Q(w,\hat{p})$ and $w = \Gamma(\hat{p}, q)$ (24) in (a.4),

$$\begin{aligned} e_u^h (u_h, \hat{p}, q) &= e_w (\Gamma(\hat{p}, q), \hat{p}, q)\, f^h (u_h, \Gamma(\hat{p}, q)) \\ &=: \alpha (\hat{p}, q)\, f^h(u_h, \Gamma(\hat{p}, q)). \end{aligned} \tag{a.5}$$

Integrating,

$$e^h(u_h, \hat{p}, q) = \alpha(\hat{p}, q)\, F^h\,(u_h, \Gamma(\hat{p}, q)) + \beta^h(\hat{p}, q) \tag{a.6}$$

and

$$e\,(w, \hat{p}, q) = \alpha(\hat{p}, q) \sum_h \left[F^h\,(u_h, \Gamma(\hat{p}, q)) + \beta^h(\hat{p}, q)\right] \tag{a.7}$$

$$= \alpha(\hat{p}, q)\, F\,(w) + \beta(p, q).$$

$F(w)$ may be replaced by w in (a.7) by using a harmless normalization. Testing (a.6) and (a.7) against (21) and (22) yields $F^h_\Gamma = 0$, and hence (29) and (30).

Case 2: In this case, (a.4) becomes

$$e^h_u\,(u_h, \bar{Q}\,(w)) = e_w\,(w, \bar{Q}\,(w))\, f^h\,(u_h, w) \tag{a.8}$$

where $\bar{Q}$ is the vector of endogenously determined prices. Since $w = \Gamma^j\,(q)$ for each j, (a.7) becomes

$$e^h(u_h, q) = \alpha(q)\, F^h\,(u_h, \Gamma^1(q), \Gamma^2(q), \ldots, \Gamma^m(q)) + \beta^h(q). \tag{a.9}$$

An analogous argument to the one above makes

$$F^h(u_h, \Gamma^1(q), \ldots, \Gamma^m(q)) = \psi^h\,(u_h). \;\square$$

Lemma 1: If $u^*_h \;\varepsilon\;$ int D^h for all h, then

$$u_h^* = - \frac{e_{a_h}(w,p,a)}{e_w(w,p,a)} \tag{59}$$

Proof: Since

$$e(w,p,a) = \sum_h e^h(u_h^*, p) \tag{a.10}$$

for all w and a,

$$e_{a_h}(w,p,a) = \sum_h e_u^h(u_h^*,p) \frac{\partial u_h^*}{\partial a_h} . \tag{a.11}$$

Since $u_h^* \, \varepsilon \, \text{int } D^h$,

$$e_u^h(u_h^*,p) = \lambda a_h \tag{a.12}$$

where λ is a Lagrange multiplier. Hence,

$$e_{a_h}(w,p,a) = \lambda \sum_h a_h \frac{\partial u_h^*}{\partial a_h} . \tag{a.13}$$

Because

$$\sum_h a_h u_h^* = w \tag{a.14}$$

for all a,

$$\Sigma a_h \frac{\partial u_h^*}{\partial a_h} + u_h^* = 0. \tag{a.15}$$

Consequently

$$e_{a_h}(w,p,a) = -\lambda u_h^* \ . \qquad (a.16)$$

By a similar argument,

$$e_w(w,p,a) = \lambda \ , \qquad (a.17)$$

and since $\lambda \neq 0$, the conclusion follows. □

Lemma 2: If $u_h^* \varepsilon$ int D^h and $a_h > 0$ for all h,

$$\frac{\partial u_h^*}{\partial w} > 0.$$

Proof: The first-order conditions for u_h^* are

$$e_u^h(u_h^*, p) = \lambda a_h \qquad (a.19)$$

where λ is a Lagrange multiplier. This holds for all w, and so

$$e_{uu}^h(u_h^*, p) \frac{\partial u_h^*}{\partial w} = \frac{\partial \lambda}{\partial w} a_h \ , \qquad (a.20)$$

$h = 1,\ldots,H$. Since e is strictly convex in u, $e_{uu}^h(u_h^*, p) > 0$. Hence the sign of $\partial u_h^*/\partial w$ must be the same for all w. Since $a_h > 0$, they must all be positive. □

Theorem 2: Given our regularity conditions, the social cost-of-living index is reference-level free for all $a > 0$ if and only if individual households' expenditure functions may be written as

Case 1:

$$e^h(u_h, p) = \Pi(p)\psi^h(u_h), \tag{61}$$

where ψ^h is a strictly convex function in $h = 1,\ldots,H$,

Case 2:

$$e^h(u_h, p) = \Pi^h(p) \exp\{u_h\} \tag{62}$$

where D^h is $\mathbb{R}$, $h = 1,\ldots,H$,

Case 3:

$$e^h(u_h, p) = \Pi^h(p)\left[r(u_h - c_h)\right]^{\frac{1}{r}}, \tag{63}$$

where $D^h = \{u_h \mid u_h \geq c_h\}$, $0 < r < 1$,

and $D^h = \{u_h \mid u_h \leq c_h\}$, $r < 0$, $h = 1,\ldots,H$.

Corresponding to these household preferences, the social cost-of-living indexes are

Case 1:

$$I(p^1, p^0, w, a) = \frac{\Pi(p^1)}{\Pi(p^0)}, \tag{64}$$

Case 2:

$$I(p^1, p^0, w, a) = \underset{h}{X} \left[\frac{\Pi^h (p^1)}{\Pi^h (p^0)}\right]^{\bar{a}_h}, \tag{65}$$

where $\bar{a}_h := a_h / \sum_g a_g$,

Case 3:

$$I(p^1, p^0, w, a) = \frac{\left[\sum_h a_h^{\frac{1}{1-r}} \Pi^h(p^1)^{\frac{-r}{1-r}}\right]^{\frac{1-r}{-r}}}{\left[\sum_h a_h^{\frac{1}{1-r}} \Pi^h(p^0)^{\frac{-r}{1-r}}\right]^{\frac{1-r}{-r}}} \tag{66}$$

Proof: Using our regularity condition, we choose $a = \bar{a} >> 0$ such that $u_h^* \in$ int D. Thus,

$$u_h^* = \bar{\alpha}^h (p, \bar{a})\, \Gamma(w, \bar{a}) + \psi^h (w, \bar{a}). \tag{60}$$

We may suppress $\bar{a}$ in (60) by writing it as

$$u_h^* = \bar{\alpha}^h (p)\, \bar{\Gamma} (w) + \psi^h (w). \tag{a.21}$$

Since $u_h^* \in$ int D^h and since

$$e^h(u_h, p) = \Pi^h (p)\, \psi^h (u_h). \tag{57}$$

$$\Pi^h (p)\, \psi^{h\prime} (u_h^*) = \lambda\, \bar{a}_h. \tag{a.22}$$

From the proof of Lemma 1,

$$e_w(w,p,a) = \lambda\,, \tag{a.17}$$

and from (53),

$$e_w(w,p,\bar{a}) = \Lambda(p,\bar{a})\, T_w(w,\bar{a}) \tag{a.23}$$

$$=: \bar{\Lambda}(p)\, k(w).$$

Using (a.17) and (a.23), (a.22) may be rewritten as

$$\frac{\psi^{h\prime}(u_h^*)}{\bar{a}_h} = \frac{\bar{\Lambda}(p)\, k(w)}{\Pi^h(p)}\,. \tag{a.24}$$

Defining $f^h(t) := \psi^{h\prime}(t)\,\bar{a}_h$, (a.24) can be rewritten, using (a.21)

$$f^h(\bar{\alpha}^h(p)\,\bar{\Gamma}(w) + \psi^h(w)) = \frac{\bar{\Lambda}(p)}{\Pi^h(p)}\, k(w). \tag{a.25}$$

By Lemma 2, we know that u_h^* is increasing in w, since e^h_{uu} is positive, f^h is increasing. Therefore, k is increasing in w.

(a.25) admits of several solutions.

Case 1: Suppose that $\bar{\alpha}^h(p)$ is constant. Then the left-hand side is independent of p, and the right-hand side must be as well. Thus, $\Pi^h(p) = c_h \bar{\Lambda}(p)$ with

$$e^h(u_h, p) = c_h\, \bar{\Lambda}(p)\, \psi^h(u_h) \tag{a.26}$$

which is equivalent to (61).

(61) implies that

$$
\begin{aligned}
e(w,p,a) &= \inf\left\{\sum_h \Pi(p)\,\psi^h(u_h) \mid \Sigma a_h u_h \geq w\right\} \\
&= \Pi(p) \inf\left\{\sum_h \psi^h(u_h) \mid \Sigma a_h u_h \geq w\right\} \qquad \text{(a.27)} \\
&= \Pi(p)\, T(w,a).
\end{aligned}
$$

Now suppose that $\bar{\alpha}^h(p)$ is not a constant and denote it as x. Then

$$
\frac{\bar{\Lambda}(p)}{\Pi^h(p)} = g^h(x), \qquad \text{(a.28)}
$$

so that (a.25) becomes

$$
f^h(x\,\bar{\Gamma}(w) + \bar{\psi}^h(w)) = g^h(x)\,k(w). \qquad \text{(a.29)}
$$

Case 3: Now suppose $\bar{\psi}^h(w) = c_h$, a constant.

Then

$$
f^h(x\,\bar{\Gamma}(w) + c_h) = g^h(x)\,k(w). \qquad \text{(a.30)}
$$

$\Gamma(w)$ must be increasing in w in this case (since u^*_h is), and we denote it as z. Then (a.30) becomes

$$
\bar{f}^h(xz) = g^h(x)\,\bar{k}(z), \qquad \text{(a.31)}
$$

a Pexider equation. x is positive, but z is not restricted in sign. The solution is (Eichhorn [1978], Blackorby and Donaldson [1982])

$$\bar{f}^h(t) = \begin{cases} \bar{b}_h \mid t \mid^{\sigma}, t \geq 0, \bar{b}_h > 0, \\ -\hat{b}_h \mid t \mid^{\sigma}, t \leq 0, \hat{b}_h > 0. \end{cases} \tag{a.32}$$

Consequently,

$$\psi^{h\prime}(t) = \begin{cases} \bar{d}_h \mid t - c_h \mid^{\sigma}, t \geq c_h, \bar{d}_h > 0, \\ -\hat{d}_h \mid t - c_h \mid^{\sigma}, t \leq \hat{c}_h, \hat{d}_h > 0. \end{cases} \tag{a.33}$$

and

$$\psi^h(u_h) = \begin{cases} \dfrac{\bar{d}_h}{\sigma+1} \mid u_h - c_h \mid^{\sigma+1}, u_h > c_h, \bar{d}_h > 0, \\ \dfrac{-d_h}{\sigma+1} \mid u_h - c_h \mid^{\sigma+1}, u_h < c_h, \hat{d}_h > 0. \end{cases} \tag{a.34}$$

Increasingness requires $\sigma + 1 \neq 0$, and strict convexity requires $(\sigma + 1) \geq 1$ for $u_h \geq c_h$, and $(\sigma + 1) \leq 1$ for $u_h \geq c_h$. However, we must have $\psi^h_{uu}(u_h) > 0$, so that $0 < (\sigma + 1) < 1$ is ruled out. Consequently, ψ^h can be written as

$$\psi^h(u_h) = d_h \left(r(u_h - c_h)\right)^{\frac{1}{r}}, \tag{a.35}$$

$0 < r < 1$ and $D^h = \{u_h \mid u_h \geq c_h\}$,

or $r < 0$ and $D^h = \{u_h \mid u_h \leq c_h\}$.

In both cases, $d_h > 0$. This yields (63), Case 3, and a simple computation gives (66).

Case 2: Now suppose that $\Gamma(w) = c$, a constant. In this case, (a.29) becomes

$$f^h(cx + \bar{\phi}^h(w)) = g^h(x)\, k(w). \tag{a.36}$$

Defining $w: = cx$ and $z: = \bar{\phi}^h (w)$ (ϕ is increasing since u^*_h is increasing), we get

$$\hat{f}^h (y + z) = \hat{g}^h (y) \hat{k} (z). \qquad \text{(a.37)}$$

This is a Pexider equation whose solution is (Eichhorn [1978])

$$\hat{f}^h(t) = c_h e^t, \qquad \text{(a.38)}$$

so that

$$\psi^{h\prime} (u_h) = \frac{c_h e^{u_h}}{a_h} \qquad \text{(a.39)}$$

and

$$\psi^h (u_h) = b_h e^{u_h}. \qquad \text{(a.40)}$$

This yields (62)(Case 2) and (65) follows from a simple computation.

We must now show that there are no other possibilities. If Γ and ϕ^h are both sensitive to w, then (a.29) holds, and we get, by differentiating,

$$f^{h\prime} (x \bar{\Gamma}(w) + \phi^h (w)) \bar{\Gamma}(w) = g^{h\prime} (x) k (w) \qquad \text{(a.41)}$$

and

$$f^{h\prime} (x \bar{\Gamma} (w) + \phi^h (w) (x \bar{\Gamma}'(w) + \phi^{h\prime} (w)) = g^h (x) k'(w). \qquad \text{(a.42)}$$

The left-hand side of (a.39) is not zero for all w, so $g'(x) \neq 0$ for some w. Further, we know $k (w) > 0$. Therefore,

$$x \frac{\bar{\Gamma}'(w)}{\bar{\Gamma}(w)} + \frac{\bar{\phi}h'(w)}{\bar{\Gamma}(w)} = \frac{g^h(x)}{g^{h'}(x)} \frac{k'(w)}{k(w)} \qquad (a.43)$$

or

$$x\hat{\Gamma}(w) + \hat{\gamma}^h(w) = \hat{g}^h(x)\hat{k}(w). \qquad (a.44)$$

If $\hat{\Gamma}(w)$ is constant, we get an immediate impossibility. If not, let $\hat{\Gamma}(w) = z$, so that (a.42) becomes

$$x\,z + \tilde{\gamma}^h(z) = \hat{g}^h(x)\,\tilde{k}(z). \qquad (a.43)$$

Differentiating,

$$z = \hat{g}^{h'}(x)\,\tilde{k}(z) \qquad (a.44)$$

requiring $g^{h'}(x) =$ a constant, $\hat{g}^h(x) = c\,x$ and $\tilde{k}(z) = z/c$. Thus, (a.43) requires $\gamma^h(z) = 0$, implying $\bar{\phi}^{h'}(w) = 0$, a contradiction. □

Footnotes

[1] We are indebted to Bert Balk, Erwin Diewert, and Margaret Slade for helpful comments and criticisms.

[2] We assume that prices are all positive, and that e^h satisfies certain regularity properties: e^h is, (i) non-negative, (ii) jointly continuous in (u_h, p), (iii) concave, positively linearly homogenous, and non-decreasing in p, (iv) continuously differentiable in (u_h, p), (v) twice continuously differentiable in u_h, (vi) increasing in u_h, and (vii) strictly convex in u_h. (vii) is not needed until Section 3.

[3] e^h and U^h must represent the same preferences. Thus, $e^h(u_h, p) = \min\limits_{x^h}\left\{\Sigma p_j x_j^h \mid U^h(x^h) \geq u_h\right\}$. See Diewert [1974].

[4] See Blackorby, Primont and Russell [1978, Lemma 3.4].

[5] See Pollak [1982].

[6] These are easily found from expenditure functions since $e\,(u, p) = y \Longleftrightarrow V\,(p, y) = u$.

[7] See Lau [1977], Jorgenson, Lau and Stoker [1981, 1982].

[8] See Diewert [1978] for a discussion.

[9] Make $\bar{u}_h = \psi^h(u_h)$ so that $\bar{e}_h(\bar{u}_h, p) := e^h(\psi^{h^{-1}}(\bar{u}_h), p)$. Since ψ^h is concave, $\psi^{h^{-1}}$ is convex, and since e^h is strictly convex in u_h, $\bar{e}^h$ is strictly convex in $\bar{u}_h$. The reasons for additive separability are mathematical tractability and the fact that we usually want indexes that apply to any number of people.

[10] Without this assumption, we would get precommitted expenditures of $\beta^h(p)$ for person h, and reference-level freedom would require $\sum\limits_h \beta^h(p) = 0$ for all p.

[11] We use a harmless additional regularity condition, namely that for all w in the interior of D (its domain in e), there exists $\bar{a} \gg 0$ such that $u_h^* \in \text{int } D^h$ for all h.

[12] called the Scitovsky-Laspeyres index. See Pollak [1980] for a discussion.

References

Blackorby, C. and D. Donaldson, ''Ratio-Scale and Translation-Scale Full Interpersonal Comparability Without Domain Restrictions: Admissible Social Evaluation Functions'', *International Economic Review*, 23, 1982, pp.249-268.

__________ , D. Primont and R. Russell, *Duality, Separability and Functional Structure: Theory and Economic Applications*, Elsevier/North-Holland, 1978.

Diewert, W.E., ''Applications of Duality Theory'', in *Frontiers of Quantitative Economics*, II, (M. Intriligator and D. Kendrick, eds.) North-Holland, 1974.

__________ , ''Hicks' Aggregation Theorem and the Existence of a Real Value-Added Function'', in *Production Economics: A Dual Approach to Theory and Applications*, Volume 2, (M. Fuss and D. McFadden, eds.), North-Holland, 1978.

Eichhorn, W., *Functional Equations in Economics*, Addison-Wesley, 1978.

Gorman, W., ''Community Preference Fields'', *Econometrica* 21, 1953, pp.63-80.

__________ , ''On a Class of Preference Fields'', *Metroeconomica*, 13, 1961, pp.53-56.

__________ , ''More Measures for Fixed Factors'', London School of Economics Mimeo, 1978.

Jorgenson, D., L. Lau and T. Stoker, ''Aggregate Consumer Behavior and Individual Welfare'', in *Macroeconomic Analysis* (D. Currie, R. Nobay and D. Peel, eds.), Croom-Helm, 1981.

__________ , ''The Transcendental Logarithmic Model of Aggregate Consumer Behavior'', in *Advances in Econometrics* V.1 (R. Baseman and G. Rhodes, eds.) JAI, 1982.

Lau, L., ''Existence of Conditions for Aggregate Demand Functions: The Case of Multiple Indexes'', Technical Report 249, IMSSS, Stanford 1977.

Muellbauer, J., ''The Political Economy of Price Indices'', Discussion Paper 22, Birkbeck College, 1974.

Nicholson, J., ''Whose Cost-of-Living?'', *Journal of the Royal Statistical Society*, Part 4, 138, 1975, pp.540-542.

Pollak, R., ''Group Cost-of-Living Indexes'', *American Economic Review*, 70, 1980, pp.273-278.

———, ''The Social Cost-of-Living Index'', *Journal of Public Economics*, 15, 1981, pp.311-336.

———, ''The Theory of the Cost-of-Living Index'', *Price Level Measurement: Proceedings From a Conference Sponsored by Statistics Canada*, 1983.

Prais, S., ''Whose Cost-of-Living?'', *Review of Economic Studies*, 26, 1959, pp.126-134.

AXIOMATIC FOUNDATION OF PRICE INDEXES AND PURCHASING POWER PARITIES

Wolfgang Eichhorn and
Joachim Voeller
Institut für Wirtschaftstheorie und Operations Research
*Universistät Karlsruhe**

SUMMARY

The paper is concerned with the so-called "axiomatic" ("statistical", "mechanistic") approach to the construction of price indexes or purchasing power parities. In this branch of index theory the theoretical foundations of both intertemporal and interspatial price comparisons are built on certain basic assumptions, called axioms, which are meant to be so general as to be satisfied by all relevant "mechanistic" price indexes/purchasing power parities. Whereas, historically, much controversy existed in the literature about the usefulness, consistency, and independence of certain sets of requirements for, generally speaking, index functions, today many propositions or theorems shed light on the various interrelations between a great number of different sets of properties for index functions. Also several characterizations of certain well-known price indexes/purchasing power parities have been published and offer an excellent chance to judge the quality of the characterized function. In this context characterization means the deduction of a certain index function from a set of given conditions such that the function in question not only meets the required conditions but also represents the only function that satisfies the conditions.

Section 1 of the paper gives a short introduction into the "axiomatic approach". Section 2 then offers an axiomatic definition of the concept of a mechanistic price index/purchasing power parity. Also, the independence of the definitional set of axioms is proven. A few well-known examples of price indexes/purchasing power parities as well as basic implications of the definitional set of axioms are presented in Section 3. Further criteria and implications are introduced in Section 4 which are partially used in Section 5 to characterize Fisher's "Ideal Index" and the so-called "Cobb-Douglas Index". It is obvious that conflicting requirements lead to inconsistent sets of conditions. Problems relating

to such inconsistencies are analysed in the final Section 6.

RÉSUMÉ

La relation entre le pouvoir d'achat de la monnaie et le niveau des prix peut être étudié soit dans l'espace soit dans le temps, soit même selon ces deux dimensions. Dans le cas de comparaisons purement temporelles, le pouvoir d'achat de la monnaie dans une certaine zone géographique, généralement un pays, est défini comme inversement proportionnel à la valeur d'un **''indice de prix''** *judicieusement choisi. Ainsi, les comparaisons temporelles de prix traitent essentiellement des problèmes méthodologiques et pratiques qui concernent la construction et l'application des indices des prix et/ou de quantité. Les comparaisons spatiales, cependant, analysent la relation entre le niveau des prix et la valeur de la monnaie à deux endroits différents à un moment donné. Bien que méthodologiquement très semblables aux comparaisons temporelles, les recherches spatiales posent des difficultés supplémentaires de théorie et de pratique. Dans toutes les comparaisons spatiales, le concept de* **''parité de pouvoir d'achat''** *est utilisé pour mesurer le niveau des prix dans un lieu comparativement à celui d'un autre lieu. Généralement, les prix sont collectés pour les biens et les services qui représentent le plus adéquatement les comportements de consommation ou de production particuliers de deux pays. Par conséquent, une parité de pouvoir d'un pays entre deux pays indique le nombre d'unités de la monnaie d'un pays qui est équivalent en pouvoir d'achat à une unité de la monnaie de l'autre pays.*

Il est évident que la parité de pouvoir d'achat peut également être calculée entre deux endroits différents d'un pays donné, par exemple, entre deux villes différentes. Dans ce cas, la parité de pouvoir d'achat compare le pouvoir d'achat de la même monnaie à deux endroits différents. De plus, il faut noter que dans certaines méthodes de comparaison entre pays, les prix respectifs de tous les pays concernés influencent la valeur d'une parité de pouvoir d'achat entre deux pays donnés.

Ce texte présente ce qu'on appelle couramment la technique ''axiomatique'' (statistique, ''mécaniste'') de construction des indices de prix ou des parités de pouvoir d'achat. Les bases théoriques de cette branche de la théorie des indices sont établies sur certains postulats de base, appelés axiomes, que l'on veut suffisamment généraux pour que tous

les indices de prix ou parités de pouvoir d'achat "mécanistes" utilisés dans les comparaisons nationales ou internationales, temporelles ou spatiales, puissent les satisfaire. Historiquement, on retrouve dans la littérature beaucoup de controverses (voir par exemple Eichhorn/Voeller [*1976*], *Voeller* [*1982*]*) sur l'utilité, la cohérence et l'indépendance de certains ensembles d'exigences pour les fonctions-indices d'une façon générale. Aujourd'hui, cependant, de nombreuses propositions ou théorèmes éclairent les diverses interrelations entre un grand nombre d'ensembles différents de propriétés pour les fonctions-indices. De plus, plusieurs caractérisations de certains indices de prix ou de parités de pouvoir d'achat bien connus ont été publiées et offrent une excellente occasion de juger de la qualité de la fonction caractérisée. Dans ce contexte, caractérisation signifie la déduction d'une certaine fonction-indice d'un ensemble de conditions données telles que la fonction en question non seulement réponde aux exigences fixées, mais constitue également la seule fonction qui y réponde.*

La Section 1 du texte est une courte introduction à la "technique axiomatique". La Section 2 présente une définition axiomatique du concept d'un indice de prix ou d'une parité de pouvoir d'achat mécaniste. On y prouve également l'indépendance d'un ensemble d'axiomes servant à la définition. Nous indiquons que les prix et les quantités sont des variables indépendantes dans notre définition. Si, cependant, on prend en considération les préférences des ménages, les prix et les quantités deviennent interdépendants. On parle alors d'indices des prix "économiques" ou d'"indices de coût de la vie". La théorie de ces indices est exposée d'une façon indépendante dans ce volume par Diewert et Pollak. Ces deux auteurs déduisent également des liens entre les indices de prix mécanistes et économiques. Dans la Section 3, on présente quelques exemples bien connus d'indices de prix ou de parités de pouvoir d'achat ainsi que les déductions fondamentales de l'ensemble d'axiomes servant de définition. Dans la Section 4, on présente des critères et des déductions supplémentaires, qu'on utilise partiellement dans la Section 5 pour décrire "l'indice idéal" de Fisher et l'indice dit "de Cobb-Douglas". Il est évident que des exigences contradictoires conduisent à des ensembles de conditions incohérentes. La Section 6, la dernière, analyse les problèmes reliés à de telles incohérences.

1. Introduction

The relationship between the purchasing power of money and the level of prices can be studied either over time or space or even with respect to both dimensions. In the case of purely **intertemporal** price comparisons, the purchasing power of money in a certain geographical area, usually a country, is defined to be inversely related to the value of some appropriately chosen **price index**. Thus, price comparisons over time essentially deal with the methodological and practical problems concerning the construction and the application of price and/or quantity indexes. **Interspatial** price comparisons, however, analyse the relationship between the price level and the value of money in two different places at a certain point of time. Though methodologically quite similar to time-to-time comparisons, place-to-place investigations pose additional difficulties both in theory and practice. In all interspatial comparisons the concept of **purchasing power parity** is used to measure the price level in one location relative to that in another. Generally, the prices are collected for those goods and services that represent most adequately particular consumption or production patterns in two countries. Therefore a purchasing power parity between two countries indicates the number of currency units of one country equivalent in purchasing power to one currency unit of the other country.

It is obvious that purchasing power parities can also be computed between two different places in a certain country, for instance, between two different cities. In this case the purchasing power parity compares the purchasing power of the same currency at two locations. Also, the fact should be noted that in certain methods for multi-country comparisons the respective prices of all countries involved influence the value of a purchasing power parity between any two countries.

The following sections set out what is commonly called the ''axiomatic'' (''statistical'', ''mechanistic'') approach to the construction of price indexes or purchasing power parities. The theoretical foundations of this branch of index theory are built on certain basic assumptions, called axioms, which are meant to be so general as to be satisfied by all relevant

price indexes/purchasing power parities in national or international comparisons over time or space. Historically, much controversy existed in the literature (see, for instance, Eichhorn/Voeller [1976], Voeller [1982]) about the usefulness, consistency and independence of certain sets of requirements for, generally speaking, index functions. Today, however, numerous propositions or theorems shed light on the various interrelations between a great number of different sets of properties for index functions. Also several characterizations of certain well-known price indexes/purchasing power parities have been published and offer an excellent chance to judge the quality of the characterized function. In this context characterization means the deduction of a certain index function from a set of given conditions such that the function in question not only meets the required conditions but also represents the only function that can be deduced from these conditions.

In closing this brief introduction we take a closer look at the contents of Sections 2-7. In Section 2 an axiomatic definition of the concept of a price index/purchasing power parity is given and the independence of the definitional set of axioms is proven. Besides a few well-known examples of price indexes/purchasing power parities Section 3 presents basic implications. Further criteria and implications are introduced in Section 4 which are partially used in Section 5 to characterize two famous indexes. It is obvious that conflicting requirements will lead to inconsistent sets of conditions, and problems relating to such inconsistencies are investigated in Section 6. Throughout the paper many abbreviated references are made to the literature which can be found in detail in the Bibliography in Section 7.

2. Notation and Definitions

It is assumed that n goods and services of identical or equivalent quality are priced either

- at two different points of time in the same place or
- in two different locations (usually two different countries) at the same point of time.

The assumption of a common list of commodities is made for simplicity reasons. Clearly, the identification of equivalent items requires

- that the time interval between the two points of time is not too long;
- that the two countries do not fall too far apart as far as their levels of economic and social development are concerned,

respectively. In both cases the selection of the commodities from the respective production or consumption pattern must not impair their representativeness, that is, the goods and services chosen should adequately represent either the situation at two different points of time or in two distinct locations (countries).

Let

$$q^A = (q_1^A,\dots,q_n^A) \in \mathbb{R}^n_{++} \qquad \left(\mathbb{R}_{++} := \{r \mid r \in \mathbb{R}, r > 0\}\right)$$

$$q^B = (q_1^B,\dots,q_n^B) \in \mathbb{R}^n_{++}$$

be the vectors of the quantities or weights of the n items purchased

- at a **base time A** and a **comparison time B** or
- at a **base country (location) A** and a **comparison country (location) B**,

respectively. The corresponding prices of these commodities are given by the price vectors

$$p^A = (p_1^A,\dots,p_n^A) \in \mathbb{R}^n_{++}$$

$$p^B = (p_1^B,\dots,p_n^B) \in \mathbb{R}^n_{++}\,.$$

Note that in the case of comparisons between countries the prices are given in different currencies. We point out here that throughout our paper the prices and quantities are **independent** variables. In this case one speaks of **mechanistic** or **statistical** price indexes. If the preferences of the households are taken into consideration, prices and quantities **depend on each other**. Then one speaks of **economic** price indexes or **cost-of-living** indexes. The theory of these indexes is independently developed in this volume by Diewert and Pollak. Both of them deduce connections between mechanistic and economic price indexes.

(2.0) **Definition:**

A function

$$P: \mathbb{R}^{4n}_{++} \to \mathbb{R}_{++}, \; (q^A, p^A, q^B, p^B) \mapsto P(q^A, p^A, q^B, p^B)$$

is called a **mechanistic (statistical) price index** (in the case of intertemporal comparisons) or a **purchasing power parity for country B with respect to country A** (in the case of interspatial comparisons) if P satisfies the following four axioms (2.1) to (2.4) for all $(q^A, p^A, q^B, p^B) \in \mathbb{R}^{4n}_{++}$. Then the value $P(q^A, p^A, q^B, p^B)$ represents the value of the price index or the purchasing power parity at the price-quantity situation (q^A, p^A, q^B, p^B).

Systems of axioms for cost-of-living indexes can be found in Diewert's, and Pollak's, contributions to this volume.

(2.1) **Monotonicity Axiom:**

The function P is strictly increasing with respect to p^B and strictly decreasing with respect to p^A:

$$P(q^A, p^A, q^B, p^B) > P(q^A, p^A, q^B, \bar{p}^B) \quad \text{if } p^B \geq \bar{p}^B,$$ [1]

$$P(q^A, p^A, q^B, p^B) < P(q^A, \bar{p}^A, q^B, p^B) \quad \text{if } p^A \geq \bar{p}^A.$$

(2.2) **Proportionality Axiom:**

If all corresponding prices differ by the same factor λ ($\lambda \in \mathbb{R}_{++}$), then the value of the function P equals λ:

$$P(q^A, p^A, q^B, \lambda p^A) = \lambda \qquad (\lambda \in \mathbb{R}_{++}).$$

(2.3) **Price Dimensionality Axiom:**

The same proportional change in the unit of the currency (currencies) does not change the value of the function P:

$$P(q^A,\lambda p^A,q^B,\lambda p^B) = P(q^A,p^A,q^B,p^B) \qquad (\lambda \in \mathbb{R}_{++}).$$

(2.4) **Commensurability Axiom:**

The same change in the units of measurement of the corresponding commodities does not change the value of the function P:

$$P(\frac{q_1^A}{\lambda_1},\ldots,\frac{q_n^A}{\lambda_n},\lambda_1 p_1^A,\ldots,\lambda_n p_n^A,\frac{q_1^B}{\lambda_1},\ldots,\frac{q_n^B}{\lambda_n},\lambda_1 p_1^B,\ldots,\lambda_n p_n^B)$$

$$= P(q^A,p^A,q^B,p^B) \qquad (\lambda_1,\ldots,\lambda_n \in \mathbb{R}_{++}).$$

Conditions (2.1) to (2.4) are called axioms since all four requirements are economically reasonable and thus represent basic properties which are desirable for every price index or purchasing power parity. Compliance with this definitional set of axioms then implies that the function P in question sensitively registers either the change of a country's price level or the purchasing power of one country's currency with respect to the other country's money.

At this point attention is called to the fact that a certain price index/purchasing power parity P is based on a given basket of goods. If, for instance, $n=1$, then $P: \mathbb{R}^4_{++} \rightarrow \mathbb{R}_{++}$ satisfying (2.1) to (2.4) is given by the so-called **price relative** or **price ratio** of the particular commodity, i.e.,

$$P(q_1^A,p_1^A,q_1^B,p_1^B) = \frac{p_1^B}{p_1^A}. \tag{2.5}$$

The proof is trivial since the Proportionality Axiom (2.2) gives

$$P(q_1^A, p_1^A, q_1^B, p_1^B) = P(q_1^A, p_1^A, q_1^B, \frac{p_1^B}{p_1^A} p_1^A) = \frac{p_1^B}{p_1^A}$$

which satisfies (2.1) to (2.4). □

Hence, in the special case of only one commodity, Axiom (2.2) is sufficient to characterize the index function. It is obvious that the shape of a function P cannot be determined so easily in the case of $n \geq 2$. In fact, the set or class of all functions satisfying (2.1) to (2.4) is not known yet. Some examples will be presented in the next section but before that, it is shown that the set of Conditions (2.1) to (2.4) is not redundant.

(2.6) Theorem:

Let $n \geq 2$. Axioms (2.1) to (2.4) are independent in the following sense: Any three of these axioms can be satisfied by a function P which does not satisfy the remaining axiom.

Proof:

The function denoted by

$$P(q^A, p^A, q^B, p^B) = \left(\frac{p_1^B}{p_1^A}\right)^{\alpha_1} \left(\frac{p_2^B}{p_2^A}\right)^{\alpha_2} \cdots \left(\frac{p_n^B}{p_n^A}\right)^{\alpha_n} \qquad (2.7)$$

$$(\alpha_1 < 0, \alpha_2 > 0, \ldots \alpha_n > 0 \text{ real constants, } \Sigma\alpha_i = 1)$$

fulfills Axioms (2.2), (2.3), (2.4), but not Axiom (2.1). The function given by

$$P(q^A, p^A, q^B, p^B) = \sum_{i=1}^{n} \frac{p_i^B}{p_i^A} \qquad (2.8)$$

satisfies Axioms (2.1), (2.3), (2.4), but not Axiom (2.2). The function represented by[2]

$$P(q^A,p^A,q^B,p^B) = \frac{q^A p^A}{q^A p^A + 1} \frac{1}{n} \sum_{i=1}^{n} \frac{p_i^B}{p_i^A} + \frac{1}{q^A p^A + 1} \max \left\{ \frac{p_1^B}{p_1^A}, \ldots, \frac{p_n^B}{p_n^A} \right\} \tag{2.9}$$

meets Axioms (2.1), (2.2), (2.4), but not Axiom (2.3). (Formula (2.9) was developed by H. Funke.) Finally the function given by

$$P(q^A,p^A,q^B,p^B) = \frac{\Sigma\, p_i^B}{\Sigma\, p_i^A} \tag{2.10}$$

conforms to Axioms (2.1), (2.2), (2.3), but not to Axiom (2.4). □

Evidently, the functions (2.7), (2.8), (2.9) and (2.10) cannot be regarded as price indexes/purchasing power parities in the sense of Definition (2.0).

3. Examples and Implications

As examples of price indexes/purchasing power parities, i.e., functions satisfying Axioms (2.1) to (2.4) the following well-known indexes[3] are presented:

$$P(q^A,p^A,q^B,p^B) = \frac{q^A p^B}{q^A p^A} \qquad \text{("Laspeyres index");} \tag{3.1}$$

$$P(q^A,p^A,q^B,p^B) = \frac{q^B p^B}{q^B p^A} \qquad \text{("Paasche index").} \tag{3.2}$$

The Laspeyres index using weights of time/country A unfortunately neglects time/country B's weights (consumption or production pattern). On the other hand, the Paasche index

using weights in B does not consider the weights in A. As a compromise the geometric mean of the two indexes was favored by Fisher [1927]:

$$P(q^A,p^A,q^B,p^B) = \left[\frac{q^A p^B}{q^A p^A}\,\frac{q^B p^B}{q^B p^A}\right]^{\frac{1}{2}} \quad \text{("Fischer's ideal index");} \tag{3.3}$$

of which

$$P(q^A,p^A,q^B,p^B) = \left(\frac{q^A p^B}{q^A p^A}\right)^{\alpha} \cdot \left(\frac{q^B p^B}{q^B p^A}\right)^{1-\alpha} \quad (0 \leq \alpha \leq 1) \tag{3.4}$$

is a generalization. The convex combination

$$P(q^A,p^A,q^B,p^B) = \alpha \frac{q^A p^B}{q^A p^A} + (1-\alpha)\frac{q^B p^B}{q^B p^A} \quad \left(0 \leq \alpha \leq 1; \text{"Drobisch index" for } \alpha = \tfrac{1}{2}\right) \tag{3.5}$$

of the indexes (3.1) and (3.2) generalizes the arithmetic mean of the Laspeyres and Paasche index.

Average weights, as applied in most of the following examples, indicate a compromise situation, too.

$$P(q^A,p^A,q^B,p^B) = \frac{(\alpha q^A + (1-\alpha)q^B)p^B}{(\alpha q^A + (1-\alpha)q^B)p^A} \quad \left(0 \leq \alpha \leq 1; \text{"Marshall-Edgeworth index" for } \alpha = \tfrac{1}{2}\right). \tag{3.6}$$

Note that the functions given by (3.4) to (3.6) represent the Laspeyres index (3.1) for $\alpha = 1$ and the Paasche index for $\alpha = 0$.

$$P(q^A,p^A,q^B,p^B) = \frac{\sum_{i=1}^{n} \sqrt{q_i^A q_i^B}\; p_i^B}{\sum_{i=1}^{n} \sqrt{q_i^A q_i^B}\; p_i^A} \qquad \text{("Walsh index");} \tag{3.7}$$

$$P(q^A,p^A,q^B,p^B) = \frac{\sum_{i=1}^{n} \frac{q_i^A\ q_i^B}{q_i^A + q_i^B}\, p_i^B}{\sum_{i=1}^{n} \frac{q_i^A\ q_i^B}{q_i^A + q_i^B}\, p_i^A} \qquad \text{("Geary-Khamis-index");} \tag{3.8}$$

$$P(q^A,p^A,q^B,p^B) = \left(\frac{p_1^B}{p_1^A}\right)^{\alpha_1} \left(\frac{p_2^B}{p_2^A}\right)^{\alpha_2} \cdots \left(\frac{p_n^B}{p_n^A}\right)^{\alpha_n} \qquad \text{("Cobb-Douglas index")} \tag{3.9}$$

$(\alpha_1 > 0,\ldots,\alpha_n > 0$ real constants, $\Sigma\alpha_i = 1)$.

For the assessment of Definition (2.0) the following theorem offers further arguments. It shows that the fulfillment of Axioms (2.1) to (2.4) leads to compliance with additional conditions which all make good economic sense. In accordance with "Irving Fisher's test approach" these conditions are called tests. Fisher [1927] used his tests to derive price index formulas and to test their quality.

(3.10) **Theorem**:

Every price index/purchasing power parity, i.e., every function P satisfying Axioms (2.1) to (2.4) also meets the following tests (3.11) to (3.14).

(3.11) **Identity Test**:

If all corresponding prices do not differ in absolute terms (but possibly in denomination), then the value of the function P equals 1:

$$P(q^A,p^A,q^B,p^A) = 1.$$

(3.12) **Weak Proportionality Test:**

If the quantity vectors satisfy $q^A = q^B$ and if all corresponding prices differ by the same factor λ ($\lambda \in \mathbb{R}_{++}$), then the value of the function P equals λ:

$$P(q^A,p^A,q^A,\lambda p^A) = \lambda. \qquad (\lambda \in \mathbb{R}_{++}).$$

(3.13) **Quantity Dimensionality Test:**

The same proportional change in all quantity units (weights) does not alter the value of the function P:

$$P(\mu q^A,p^A,\mu q^B,p^B) = P(q^A,p^A,q^B,p^B) \qquad (\mu \in \mathbb{R}_{++}).$$

(3.14) **Mean Value Test:**

The value of the function P always lies between the smallest and the largest price ratio of the corresponding prices:

$$\min\left\{\frac{p_1^B}{p_1^A},\ldots,\frac{p_n^B}{p_n^A}\right\} \le P(q^A,p^A,q^B,p^B) \le \max\left\{\frac{p_1^B}{p_1^A},\ldots,\frac{p_n^B}{p_n^A}\right\}.$$

Proof of Theorem (3.10):

The Identity Test (3.11) and the Weak Proportionality Test (3.12) immediately follow from the Proportionality Axiom (2.2). To prove the assertion in the case of (3.13), Axiom (2.3) implies

$$P(\mu q^A,p^A,\mu q^B,p^B) = P(\mu q^A,\lambda p^A,\mu q^B,\lambda p^B).$$

After setting $\mu = \frac{1}{\lambda}$ and applying the Commensurability Axiom (2.4) for $\lambda_1 = \lambda_2 = \ldots = \lambda_n = \lambda$

$$P\left(\frac{q^A}{\lambda}, \lambda p^A, \frac{q^B}{\lambda}, \lambda p^B\right) = P(q^A, p^A, q^B, p^B)$$

is derived.

All that remains is to prove the proposition for the Mean Value Test. We first define

$$a := \min\left\{\frac{p_1^B}{p_1^A}, \ldots, \frac{p_n^B}{p_n^A}\right\} \text{ and } b := \max\left\{\frac{p_1^B}{p_1^A}, \ldots, \frac{p_n^B}{p_n^A}\right\}.$$

Now, on one hand

$$\underset{\text{(by (2.2))}}{a = P(q^A, p^A, q^B, ap^A)} \underset{\text{(by (2.1))}}{\leq} P(q^A, p^A, q^B, p^B)$$

and, on the other hand

$$\underset{\text{(by (2.2))}}{b = P(q^A, p^A, q^B, bp^A)} \underset{\text{(by (2.1))}}{\geq} P(q^A, p^A, q^B, p^B).$$

Hence,

$$\min\left\{\frac{p_1^B}{p_1^A}, \ldots, \frac{p_n^B}{p_n^A}\right\} \leq P(q^A, p^A, q^B, p^B) \leq \max\left\{\frac{p_1^B}{p_1^A}, \ldots, \frac{p_n^B}{p_n^A}\right\}. \square$$

Methods for constructing new price indexes or purchasing power parities from given ones are described in the following three theorems. The proof of Theorem (3.15) is obvious and the proofs for Theorems (3.18) and (3.20) can be found in Krtscha [1979].

(3.15) Theorem:

If $P_1,\ldots,P_k$ are price indexes/purchasing power parities in accordance with Definition (2.0), then

$$(\alpha_1 P_1^\delta + \ldots + \alpha_k P_k^\delta)^{1/\delta} \qquad \begin{cases} \delta \neq 0,\ \alpha_1 \geq 0, \ldots, \alpha_n \geq 0 \\ \text{real constants, } \Sigma\alpha_i = 1 \end{cases} \tag{3.16}$$

and

$$P_1^{\delta_1} \cdot P_2^{\delta_2} \cdot \ldots \cdot P_k^{\delta_k} \qquad \begin{cases} \delta_1 \geq 0, \ldots, \delta_k \geq 0, \\ \text{real constants, } \Sigma\delta_i = 1 \end{cases} \tag{3.17}$$

also represent price indexes/purchasing power parities. P^δ is defined as

$$(q^A,p^A,q^B,p^B) \mapsto [P(q^A,p^A,q^B,p^B)]^\delta.$$

(3.18) Theorem:

If $P_1,\ldots,P_k$ are price indexes/purchasing power parities in accordance with Definition (2.0), then the function defined by

$$(q^A,p^A,q^B,p^B) \mapsto \bar{P}(P_1(q^A,p^A,q^B,p^B),\ldots,P_k(q^A,p^A,q^B,p^B)) \tag{3.19}$$

is also a price index/purchasing power parity if $\bar{P}: \mathbb{R}^k_{++} \to \mathbb{R}_{++}$ is linearly homogenous, monotonically increasing and $\bar{P}(1,\ldots 1) = 1$.

(3.20) **Theorem:**

If P is a price index/purchasing power parity according to Definition (2.0), then

$$P(a_1q_1^A,\ldots,a_nq_n^A,b_1p_1^A,\ldots,b_np_n^A,c_1q_1^B,\ldots,c_nq_n^B,b_1p_1^B,\ldots,b_np_n^B) \tag{3.21}$$

$(a_1,\ldots,a_n,b_1,\ldots,b_n,c_1,\ldots,c_n$ positive real constants)

also represents a price index/purchasing power parity.

We point out here that a procedure or a set of procedures that generates **all** functions P satisfying Axioms (2.1) to (2.4), i.e., the set of all mechanistic (statistical) price indexes, is not yet known.

It was the objective of this section to throw some light on the reasonableness of Definition (2.0). It must be admitted, however, that the class of functions obeying (2.0) is still extremely wide, for an infinite number of functions P: $\mathbb{R}^{4n}_{++} \to \mathbb{R}_{++}$ can also be generated according to Theorems (3.15), (3.18), (3.20) and other procedures. Therefore, in the next section we go one step further and present additional conditions (tests, criteria) which, required in addition to Axioms (2.1) to (2.4), may drastically reduce the number of functions conforming to the new set of conditions. Later, in Section 5, we will look at those sets of axioms/tests where only one particular price index/purchasing power parity or just one class of functions P satisfies the respective set of requirements.

4. Further Criteria and Implications

Most of the requirements now mentioned are again stated as functional equations which are valid for all (q^A,p^A,q^B,p^B):

(4.1) **Linear Homogeneity Test:**

If all comparison time/country prices change λ-fold ($\lambda \in \mathbb{R}_{++}$), then the value of the

function P is changed by λ:

$$P(q^A, p^A, q^B, \lambda p^B) = \lambda P(q^A, p^A, q^B, p^B) \qquad (\lambda \in \mathbb{R}_{++}).$$

(4.2) **Theorem**:

Every function $P: \mathbb{R}^{4n}_{++} \to \mathbb{R}_{++}$ satisfying the Linear Homogeneity Test (4.1) and the Identity Test (3.11) also fulfills the Proportionality Axiom (2.2). The converse is not true.

Proof:

After setting $p^B = p^A$, Tests (4.1) and (3.11) imply the Proportionality Axiom (2.2). That the converse does not hold is verified by the following function:

$$P(q^A, p^A, q^B, p^B) = \frac{q^A(p^A + p^B)}{q^A p^A} \cdot \frac{q^B p^B}{q^B(p^A + p^B)} \tag{4.3}$$

which meets the Proportionality Axiom (2.2) and, hence, the Identity Test (3.11), but not the Linear Homogeneity Test (4.1). □

In addition, the function given by (4.3) also satisfies the Monotonicity Axiom (2.1), the Price Dimensionality Axiom (2.3), and the Commensurability Axiom (2.4). Hence, Corollary (4.4) is immediately deduced:

(4.4) **Corollary**:

The class of functions $P: \mathbb{R}^{4n}_{++} \to \mathbb{R}_{++}$ conforming to Axioms (2.1) to (2.4) is wider than the class of functions satisfiying Axioms (2.1), (2.3), (2.4), the Identity Test (3.11) and the Linear Homogeneity Test (4.1).

In our monograph (Eichhorn/Voeller [1976]; see also Eichhorn [1976, 1978]) we used Axioms (2.1), (2.3), (2.4) and Tests (3.11), (4.1) to define price indexes.

Next, four tests are considered three of which impose a kind of reversal procedure on a function $P: \mathbb{R}^{4n}_{++} \rightarrow \mathbb{R}_{++}$. Three of the tests, that is, the Time/Country Reversal Test, the Factor Reversal Test and the Product Test (Weak Factor Reversal Test) originally date back to I. Fisher's investigations [1927] while the Price Reversal Test was first published in Funke/Voeller [1979]. All four tests deserve attention since they will be used to characterize a certain price index/purchasing power parity.

(4.5) **Time/Country Reversal Test**:

If $P(q^A,p^A,q^B,p^B)$ compares quantities and prices of comparison time/country B with those of base time/country A and $P(q^B,p^B,q^A,p^A)$ does the same in the opposite direction, then the product of the values of the two function values equals one:

$$P(q^A,p^A,q^B,p^B)\cdot P(q^B,p^B,q^A,p^A) = 1.$$

The fulfillment of this test implies that it does not matter which time/country is taken as denominator time/country (''base time/country invariance''). One index P can be easily calculated as the reciprocal of the other. Kravis *et al.* [1975, p.47] comment on this requirement, referring only to interspatial comparisons, as follows: ''In many applications the utility of the results would be greatly diminished if, for example, it were necessary to provide two estimates of the comparison between the countries depending upon which was taken as the denominator country in the ratio between them.''

(4.6) **Factor Reversal Test**:

In P, q^A and p^A as well as q^B and p^B are interchanged. The resulting $P(p^A,q^A,p^B,q^B)$ can be regarded as the value of a quantity index if $P(q^A,p^A,q^B,p^B)$ is the value of a price index/purchasing power parity.[4] The product of the two values equals the expenditure ratio (value ratio) for the two baskets of goods:

$$P(q^A,p^A,q^B,p^B)\cdot P(p^A,q^A,p^B,q^B) = \frac{q^Bp^B}{q^Ap^A}.$$

A weaker version of Test (4.6) represents the Product Test:

(4.7) **Product Test:**

The product of the values of P and of a quantity index Q: $\mathbb{R}^{4n}_{++} \rightarrow \mathbb{R}_{++}$ which satisfies axioms analogous to (2.1) to (2.4) equals the expenditure ratio:

$$P(q^A,p^A,q^B,p^B)\cdot Q(q^A,p^A,q^B,p^B) = \frac{q^Bp^B}{q^Ap^A}.$$

On the practical use of the Product Test Drechsler [1973, p.22] may be quoted: ''One should not forget here that in many cases quantity indexes cannot be compiled directly but only by deflation of value data by means of price indexes.'' Contrary to this procedure the price index/purchasing power parity and the quantity index must be computed independently in order that the stronger Factor Reversal Test (4.6) is satisfied.

A rather controversial test is given by

(4.8) **Price Reversal Test:**

If all base and comparison time/country prices are interchanged but the quantities remain unchanged, then the product of the values of the two function values equals one:

$$P(q^A,p^A,q^B,p^B)\cdot P(q^A,p^B,q^B,p^A) = 1.$$

For critical comments on this condition which, by the way, is satisfied by such famous indexes like the Laspeyres index (3.1), the Paasche index (3.2), and Fisher's ''ideal index'' (3.3) see Sato [1980] and Funke/Voeller [1980].

The next requirement contains a rather weak continuity assumption:

(4.9) **Determinateness Test:**

If any scalar argument in P tends to zero, then the value of the function P tends to a unique positive real number.

Now, two new criteria are introduced.

$$P(q^A,p^A,q^B,\lambda_1 p_1^A,\ldots,\lambda_n p_n^A) \tag{4.10}$$

$$= P(q^A,\lambda_1 p_1^A,\ldots,\lambda_n p_n^A,q^B,\lambda_1^2 p_1^A,\ldots,\lambda_n^2 p_n^A)$$

$$(\lambda_i \in \mathbb{R}_{++},\ i=1,\ldots,n).$$

In words, Condition (4.10) says:

As long as the price of each commodity in the base and comparison time/country differs by the same factor λ_i ($\lambda_i \in \mathbb{R}_{++}$, $i=1,\ldots,n$), the value of the function P remains the same.

A somewhat stronger version of Condition (4.10), is given by

$$P(q^A,p^A,q^B,\lambda_1 p_1^A,\ldots,\lambda_n p_n^A) = P(q^A,p^C,q^B,\lambda_1 p_1^C,\ldots,\lambda_n p_n^C) \tag{4.11}$$

$$(\lambda_i \in \mathbb{R}_{++},\ i=1,\ldots,n).$$

Incidentally, Conditions (4.10) and (4.11) imply that the weight that each commodity is given in the basket of goods does not change in the course of further comparisons. The often used indexes denoted by (3.1), (3.2), (3.3) and many other price indexes/purchasing power parities do not conform to this requirement while the Cobb-Douglas Index (3.9) does. A detailed description of the implications of Property (4.11) taken together with other conditions is postponed until Section 5. At this point just one notable consequence of Condition (4.11) is evaluated:

(4.12) **Remark:**

Conformance to Condition (4.11) by a function P implies that with respect to the prices

the function P only depends on the price relatives for the n commodities.

Proof:

Applying (4.11),

$$P(q^A,p^A,q^B,p^B) = P\left(q^A,p^A,q^B, \frac{p_1^B}{p_n^A}\, p_1^A,\ldots,\frac{p_n^B}{p_n^A}\, p_n^A\right)$$

$$= P\left(q^A,1,\ldots,1,q^B, \frac{p_1^B}{p_1^A},\ldots,\frac{p_n^B}{p_n^A}\right). \square$$

The presentation of tests or conditions used for intertemporal or interregional comparison methods would be incomplete if the famous Circular Test as well as its implications were not considered thoroughly. In his pathbreaking book ‘‘The Making of Index Numbers’’ I. Fisher commented on the ‘‘merit’’ of the Circular Test by saying [1927, p.271]: ‘‘I aim to show that the Circular Test is theoretically a mistaken one, that a irreducible minimum of divergence from such fulfillment is entirely right and proper, and, therefore, that a perfect fulfillment of this so-called Circular Test should really be taken as proof that the formula which fulfills it is erroneous.’’ It is not surprising that Fisher's rigorous statement has stirred up many controversies among statisticians. By and large they objected to Fisher's conclusion and generally considered the failure of any purchasing power parity (or, equivalently, any price index formula) to pass the Circular Test as a shortcoming of this particular index. As a result, the Circular Test is given credit in almost all index-theoretic publications. Especially the further development and increased application of multilateral purchasing power parities have generated a number of statistical methods satisfying the Circular Test. Thus, a clash of opinion with Fisher's assertion seems to be inevitable. A concluding answer to his contention will not be given until Section 5.

(4.13) **Circular Test:**

If (q^A,p^A,q^B,p^B) represents the value of a price index/purchasing power parity between times/locations A and B and $P(q^B,p^B,q^C,p^C)$ represents the value for a comparison between times/locations B and C, then the product of the two values of P equals the value of a price index/purchasing power parity between times/places A and C:

$$P(q^A,p^A,q^B,p^B)\cdot P(q^B,p^B,q^C,p^C) = P(q^A,p^A,q^C,p^C).$$

In many publications the Circular Test (4.13) is also called the "chain method" since the separate time-to-time or place-to-place comparisons are joined together by successive multiplications like the links of a chain. Sometimes, the term "transitivity" requirement is used, too.

A weakened version of the Circular Test is given by the Base Test whose economic meaning is essentially the same as that of (4.13). Formally, the requirements are weaker, however:

(4.14) **Base Test:**

There exist functions P and $\tilde{P}$ such that either

$$P(q^A,p^A,q^B,p^B)\cdot \tilde{P}(q^B,p^B,q^C,p^C) = P(q^A,p^A,q^C,p^C)$$

or

$$\tilde{P}(q^A,p^A,q^B,p^B)\cdot P(q^B,p^B,q^C,p^C) = P(q^A,p^A,q^C,p^C)$$

holds.

Note that $\widetilde{P}$ only depends on four of the six vectors occurring in (4.14).

Now it is trivial to prove

(4.15) **Remark:**

Every function $P: \mathbb{R}^{4n}_{++} \to \mathbb{R}_{++}$ satisfying the Identity Test (3.11) and the Circular Test (4.13) also satisfies the Time/Country Reversal Test (4.5).

The Circular Test (4.13) and the Time/Country Reversal Test (4.5) lead to the following version of the Circular Test where over a number of intermediate steps a closed circuit is achieved:

$$P(q^A,p^A,q^B,p^B)\cdot P(q^B,p^B,q^C,p^C)\cdot \ldots \cdot P(q^K,p^K,q^A,p^A) = 1. \qquad (4.16)$$

From Funke/Hacker/Voeller [1979, p.682] the following theorem is taken:

(4.17) **Theorem:**

Every function $P: \mathbb{R}^{4n}_{++} \to \mathbb{R}_{++}$ satisfying the Linear Homogeneity Test (4.1) and the Circular Test (4.13) also conforms to the Price Dimensionality Axiom (2.3). The converse is not true.

Without proof as well the next theorem is cited from Eichhorn/Voeller [1976, p.33]:

(4.18) **Theorem:**

Every function $P: \mathbb{R}^{4n}_{++} \to \mathbb{R}_{++}$ fulfilling the weak Proportionality Test (3.12) and the Base Test (4.14) also satisfies the Circular Test (4.13). The converse is not true.

Finally, two variations of the Circular Test (4.13) are presented. Since the economic meaning of each condition can easily be grasped we just write down the respective functional equations:

(4.19) **Price Circular Test:**

$$P(q^X,p^A,q^Y,p^B)\cdot P(q^X,p^B,q^Y,p^C) = P(q^X,p^A,q^Y,p^C).$$

A special case of both the Circular Test (4.13) and the Price Circular Test (4.19) represents

(4.20) **Weak Circular Test:**

$$P(q^A,p^A,q^A,p^B)\cdot P(q^A,p^B,q^A,p^C) = P(q^A,p^A,q^A,p^C).$$

Now it is easy to confirm

(4.21) **Remark:**

Every function $P:\mathbb{R}^{4n}_{++} \rightarrow \mathbb{R}_{++}$ conforming to the Price Circular Test (4.19) also conforms to the Identity Test (3.11).

The following theorem, independently derived by Hacker [1979, p.77] and Krtscha [1979, p.70], provides important insight into the influence of the Circular Test (4.13) on the shape of an index function:

(4.22) **Theorem:**

Every function $P: \mathbb{R}^{4n}_{++} \rightarrow \mathbb{R}_{++}$ meeting the Identity Test (3.11) and the Circular Test (4.13) only depends on the base and comparison time/country prices but not on the quantities of the commodities.

Proof:

From (4.13) we obtain

$$P(q^A,p^A,q^B,p^B) = \frac{P(q^A,p^A,q^C,p^C)}{P(q^B,p^B,q^C,p^C)}.$$

Setting $q^C = (1,\ldots,1)$ and $p^C = (1,\ldots,1)$ it follows that

$$P(q^A,p^A,q^B,p^B) = \frac{f(q^A,p^A)}{f(q^B,p^B)}, \quad \text{where } f(q^A,p^A) = P(q^A,p^A,1,\ldots,1).$$

Applying the Identity Test (3.11) yields

$$P(q^A,p^A,q^B,p^A) = \frac{f(q^A,p^A)}{f(q^B,p^A)} = 1,$$

hence,

$$f(q^A,p^A) = f(q^B,p^A).$$

Obviously f does not depend on q^A or q^B. Therefore we have

$$f(q^A,p^A) = g(p^A)$$

which leads to

$$P(q^A,p^A,q^B,p^B) = \frac{g(p^A)}{g(p^B)}. \quad \square$$

(4.23) **Corollary**:

If the Circular Test (4.13) is replaced by the Base Test (4.14) in Theorem (4.22), then the same implication holds.

The **proof** runs analogously to the proof of Theorem (4.22).

Since the Identity Test (3.11) is a special case of the Proportionality Axiom (2.2) the following corollary can be stated:

(4.24) **Corollary**:

Every function P: $\mathbb{R}^{4n}_{++} \to \mathbb{R}_{++}$ satisfying the Proportionality Axiom (2.2) and the Circular Test (4.13) (or the Base Test (4.14)) only depends on the prices.

Now it is extremely important to realize that Corollary (4.24) has severe consequences for the evaluation of several so-called multilateral purchasing power parities and also for the so-called chain methods in price index theory. **Since the fulfillment of the Circular Test (4.13) is the basic feature of all those methods and since almost all use quantity weights in their calculations they cannot conform to the Proportionality Axiom (2.2).** We point out here that this non-fulfillment represents a decisive drawback. Note that such an index cannot be regarded as a price index/purchasing power parity according to Definition (2.0).

5. Characterizations

The purpose of this section is to present characterizations of well-known price indexes/purchasing power parities. The search for characterizations is attributable not only to the intellectual challenge such a deduction represents but also to its great practical importance. For the statistician who is interested in knowing the class of functions satisfying a certain subset of requirements a characterization offers a definite answer. Characterizations of **cost-of-living indexes** are given by Diewert and Pollak in this volume.

Of course any given set of conditions should be independent in the following sense: no condition is deductible from the remaining ones. Hence, the given set of conditions is also **strict** (or minimal) which means that after dropping any one of the conditions the characterization and therefore the implications do not hold any more. Thus, redundant sets of conditions, i.e., sets containing conditions which are irrelevant to a characterization are not considered.

(5.1) **Theorem**:

The Monotonicity Axiom (2.1), the Proportionality Axiom (2.2), the Commensurability Axiom (2.4), and the Circular Test (4.13) (or the Base Test (4.14)) are independent in the

following sense: Any three conditions can be satisfied by a function $P: \mathbb{R}^{4n}_{++} \to \mathbb{R}_{++}$ which does not satisfy the remaining condition.

Proof:

The function represented by (2.7) satisfies (2.2), (2.4) and (4.13) (hence (4.14)), but not (2.1). The function given by

$$P(q^A,p^A,q^B,p^B) = \sqrt{\frac{q^Bp^B}{q^Ap^A}} \qquad (5.2)$$

fulfills (2.1), (2.4) and (4.13) (hence (4.14)), but not (2.2).

The function given by

$$P(q^A,p^A,q^B,p^B) = \frac{cp^B}{cp^A} \quad \text{(``Lowe index'',} \ c_1>0,\dots,c_n>0 \text{ real constants)} \qquad (5.3)$$

passes (2.1), (2.2) and (4.13) (hence (4.14)), but not (2.4). Finally, the function denoted by (3.1) conforms to (2.1), (2.2) and (2.4), but not to (4.13) (and not to (4.14)). □

(5.4) **Theorem**:

A function $P: \mathbb{R}^{4n}_{++} \to \mathbb{R}_{++}$ fulfills the Monotonicity Axiom (2.1), the Proportionality Axiom (2.2), the Commensurability Axiom (2.4) and the Base Test (4.14) if and only if P is given by the Cobb-Douglas Index (3.9).

Proof (see Funke/Hacker/Voeller [1979]):

"<=": This direction is easily verified by inserting (3.9) into (2.1), (2.2), (2.4) and (4.14).

"=>": As proven in Eichhorn/Voeller [1976, p.35] a function P satisfying the Commensurability Axiom (2.4) and the Base Test (4.14) can be written in the following form:

$$P(q^A,p^A,q^B,p^B) = \frac{G(q_1^B p_1^B,\dots,q_n^B p_n^B)}{H(q_1^A p_1^A,\dots,q_n^A p_n^A)}\,\Phi\left(\frac{p_1^B}{p_1^A},\dots,\frac{p_n^B}{p_n^A}\right) \tag{5.5}$$

with functions $G,H,\Phi:\mathbb{R}^n_{++} \to \mathbb{R}_{++}$ and function Φ multiplicative, i.e.,

$$\Phi(\lambda_1\mu_1,\dots,\lambda_n\mu_n) = \Phi(\lambda_1,\dots,\lambda_n)\Phi(\mu_1,\dots,\mu_n). \tag{5.6}$$

Applying the Identity Test (3.11) as a special case of the Proportionality Axiom (2.2) to (5.5) one at once derives

$$\frac{G(q_1^B p_1^A,\dots,q_n^B p_n^A)}{H(q_1^A p_1^A,\dots,q_n^A p_n^A)}\,\Phi(1,\dots,1) = 1. \tag{5.7}$$

Setting first $p^A = (1,\dots,1)$ and $q^A = (1,\dots,1)$ and then $p^A = (1,\dots,1)$ and $q^B = (1,\dots,1)$, Equation (5.7) changes to

$$G(q_1^B,\dots,q_n^B) = c \in \mathbb{R}_{++}$$

and

$$H(q_1^A,\dots,q_n^A) = c \in \mathbb{R}_{++},$$

respectively. Hence

$$P(q^A,p^A,q^B,p^B) = \Phi\left(\frac{p_1^B}{p_1^A},\dots,\frac{p_n^B}{p_n^A}\right).$$

Because of the Monotonicity Axiom (2.1) Φ is strictly increasing. This result and (5.6) imply (see Eichhorn [1978, p.66]) that Φ can be written as

$$\Phi\left(\frac{p_1^B}{p_1^A},\ldots,\frac{p_n^B}{p_n^A}\right) = \left(\frac{p_1^B}{p_1^A}\right)^{\alpha_1} \cdots \left(\frac{p_n^B}{p_n^A}\right)^{\alpha_n}$$

where $\alpha_1,\ldots,\alpha_n$ are positive real constants. The application of the Proportionality Axiom (2.2) finally yields $\Sigma\alpha_i = 1$, i.e., Formula (3.9). □

(5.8) **Remark**:

If in Theorem (5.4) the Base Test (4.14) is replaced by the Circular Test (4.13), the same result as in Theorem (5.4) holds.

The **proof** runs analogously to that of Theorem (5.4).

With Remark (5.8) the ground is laid for a conclusive answer to the old controversy about the value of the Circular Test. On page 431, Irving Fisher's standpoint was briefly indicated and it is summed up again in the following sentence [1927, p.276]: "It is clear that constant weighting, though it makes it possible to fulfill the circular test, does so at the expense of forcing the facts, for the true weights are **not** constants."

What Fisher assumed without final proof, that is, that formulas which satisfy the Circular Test are certain index numbers with constant weights is now confirmed by the results of Theorem (5.4) and Remark (5.8). In both statements the α_is of function (3.9) can be interpreted as constant weights. Besides the Cobb-Douglas Index, Fisher also noticed Lowe's Index given by (5.3) as conforming to his Circular Test. But according to Definition (2.0) Lowe's Index is not a price index/purchasing power parity since it does not fulfill the Commensurability Axiom (2.4). Thus, the fact cannot be denied that the main purpose of the Circular Test, that is, the adjustment of the quantity weights to the new situation in each new dual comparison around a circle of points of time or countries cannot be accomplished. There simply does not exist a function P which satisfies our four basic axioms for price indexes/purchasing power parities and the Circular Test, simultaneously, **and** which enables

adequately changing weights. In fact, only the Cobb-Douglas function (3.9) with its constant weighting scheme conforms to the basic requirements and the Circular Test which currently seems to be a *conditio sine qua non* specially in multilateral purchasing power parity comparisons.

Thus, the following remark can be taken as a late justification of I. Fisher's repudiation of the Circular Test:

(5.9) **Remark**:

The fulfillment of the Circular Test (4.13) or the Base Test (4.14) in all possible methods of intertemporal or interspatial price comparisons is achieved by:

(i) constant weighting schemes which contradict the ''spirit'' of the Circular or Base Test;

(ii) non-fulfillment of one or more of the four basic axioms given in Definition (2.0).

The conclusion reached in Remark (5.9) is the reason why Funke/Hacker/Voeller in [1979, p.686] proposed to discard the Circular Test and the Base Test and to use instead the two variations (4.19) and (4.20). In other words, in a circular price comparison it makes much more sense to do without the requirement of changing weights which cannot be met anyway.

A further characterization of the Cobb-Douglas Index is possible using the following requirements:

(5.10) **Theorem**:

The Monotonicity Axiom (2.1), the Proportionality Axiom (2.2), Condition (4.11), and the Circular Test (4.13) are independent in the following sense: Any three of these conditions can be satisfied by a function $P\colon \mathrm{IR}^{4n}_{++} \to \mathrm{IR}_{++}$ which does not conform to the remaining condition.

i.e., only depends on the prices. Because of Remark (4.12), Condition (4.11) implies

$$P(q^A,p^A,q^B,p^B) = P(q^A,1,\ldots,1,q^B,\frac{p_1^B}{p_1^A},\ldots,\frac{p_n^B}{p_n^A}).$$

Hence,

$$\tilde{P}\,(\frac{p_1^B}{p_1^A},\ldots,\frac{p_n^B}{p_n^A}) = \frac{\tilde{P}(p_1^B,\ldots,p_n^B)}{\tilde{P}(p_1^A,\ldots,p_n^A)}, \tag{5.14}$$

$$\text{where } \tilde{P}(\gamma) := \frac{f(\gamma)}{f(1,\ldots,1)}$$

After substituting

$$p_i^B = p_i^B p_i^A$$

we obtain

$$\tilde{P}(p_1^B,\ldots,p_n^B)\tilde{P}(p_1^A,\ldots,p_n^A) = \tilde{P}(p_1^B\,p_1^A,\ldots,p_n^B\,p_n^A), \tag{5.15}$$

i.e., a multiplicative function of type (5.6). Because of the Monotonicity Axiom (2.1) $\tilde{P}$ is strictly increasing. This result and (5.15) imply (see Eichhorn [1978, p.66])

$$\tilde{P}\,(\frac{p_1^B}{p_1^A},\ldots,\frac{p_n^B}{p_n^A}) = \left(\frac{p_1^B}{p_1^A}\right)^{\alpha_1}\left(\frac{p_2^B}{p_2^A}\right)^{\alpha_2}\cdots\left(\frac{p_n^B}{p_n^A}\right)^{\alpha_n}$$

with $\alpha_1 > 0,\ldots,\alpha_n > 0$ positive real constants. Now the application of the Proportionality Axiom (2.1) implies $\Sigma\alpha_i = 1$ and, thus, leads, to the Cobb-Douglas Index (3.9). The other direction of the proof is obvious. □

Proof:

The function given by (2.7) meets (2.2), (4.11) and (4.13), but not (2.1). The function represented by

$$P(q^A,p^A,q^B,p^B) = \left(\frac{p_1^B}{p_1^A}\right)^{\alpha_1} \left(\frac{p_2^B}{p_2^A}\right)^{\alpha_2} \cdots \left(\frac{p_n^B}{p_n^A}\right)^{\alpha_n} \tag{5.11}$$

($\alpha_1>0,\ldots,\alpha_n>0$ positive real constants, $\Sigma\alpha_i \neq 1$)

satisfies (2.1), (4.11) and (4.13), but not (2.2). The function denoted by (2.10) fulfills (2.1), (2.2) and (4.13), but not (4.11). At last, the function given by

$$P(q^A,p^A,q^B,p^B) = \frac{1}{n}\Sigma \frac{p_i^B}{p_i^A} \tag{5.12}$$

conforms to (2.1), (2.2) and (4.11), but not to (4.13). □

(5.13) **Theorem:**

A function $P\colon \mathbb{R}^{4n}_{++} \rightarrow \mathbb{R}_{++}$ satisfies the Monotonicity Axiom (2.1), the Proportionality Axiom (2.2), Condition (4.11) and the Circular Test (4.13) if and only if P is the Cobb-Douglas Index (3.9).

Proof:

According to Remark (4.24) every function P conforming to the Proportionality Axiom (2.2) and the Circular Test (4.13) can be written as

$$P(q^A,p^A,q^B,p^B) = \frac{f(p^B)}{f(p^A)},$$

Many alternative aggregation methods have been favored by one statistical agency or another but only very few have been as widely used as Irving Fisher's so-called "ideal index" denoted by (3.3). It produces results that in every instance are intermediate between its two component indexes (3.1) and (3.2). Hence, it has been regarded as an evenhanded compromise that is often applied for pragmatic reasons: it serves the interests of brevity and it avoids the cumbersome presentation of all comparisons with the two more basic indexes.

Historically, Fisher [1927, p.220 ff.] called the function given by (3.3) "ideal" since it satisfied all of his tests for price indexes except for the Circular Test (4.13). As mentioned before, Fisher therefore proposed to discard the Circular Test. However, his assumed superiority of the "ideal index" over all other index numbers has often been challenged in the literature and, today, a more application-oriented approach seems to be predominant. A particular price index/purchasing power parity is chosen with the properties desirable for a special assignment. Characterizations which offer clear-cut answers to such problems therefore deserve increased attention.

The following characterization of the Fisher Index has been originally published in Funke/Voeller [1979]. To achieve it the three Reversal Tests (4.5), (4.6) and (4.8) are sufficient.

(5.16) **Theorem**:

The Time/Country Reversal Test (4.5), the Factor Reversal Test (4.6) and the Price Reversal Test (4.8) are independent in the following sense: any two of these conditions can be satisfied by a function P: $\mathbb{R}^{4n}_{++} \to \mathbb{R}_{++}$ which does not satisfy the remaining condition.

Proof:

The function given by

$$P(q^A,p^A,q^B,p^B) = \tag{5.16}$$

$$\left[\frac{q^Ap^B}{q^Ap^A}\frac{q^Bp^B}{q^Bp^A}\right]^{\frac{1}{2}} \frac{(q^B+p^B)p^A}{(q^A+q^B)p^A}\frac{(q^B+p^A)q^A}{(q^B+p^B)q^A}\frac{(q^A+p^A)p^B}{(q^B+p^A)p^B}\frac{(q^A+p^B)q^B}{(q^A+p^A)q^B}$$

meets (4.6) and (4.8), but not (4.5). The function denoted by (3.7) conforms to (4.5) and (4.8), but not to (4.6). Finally, the function represented by (5.2) fulfills (4.5) and (4.6), but not (4.8). □

(5.17) **Theorem**:

A function $P: \mathbb{R}^{4n}_{++} \to \mathbb{R}_{++}$ satisfies the Time/Country Reversal Test (4.5), the Factor Reversal Test (4.6) and the Price Reversal Test (4.8) if and only if P is Fisher's ''ideal index'' (3.3).

Proof:

''<='': is obvious.

''=>'': From (4.5)

$$P(q^B,p^B,q^A,p^A) = \frac{1}{P(q^A,p^A,q^B,p^B)}$$

and from (4.8)

$$P(q^A,p^B,q^B,p^A) = \frac{1}{P(q^A,p^A,q^B,p^B)}$$

the following equation follows:

$$P(q^A,p^B,q^B,p^A) = P(q^B,p^B,q^A,p^A) \quad 5 \tag{5.18}$$

Interchanging p^A and p^B in (4.6) yields

$$P(q^A,p^B,q^B,p^A)\cdot P(p^B,q^A,p^A,q^B) = \frac{q^B p^A}{q^A p^B}. \tag{5.19}$$

Using (5.18), Equation (5.19) changes into

$$P(q^B,p^B,q^A,p^A)\cdot P(p^A,q^A,p^B,q^B) = \frac{q^B p^A}{q^A p^B}. \tag{5.20}$$

Dividing (4.6) by (5.20) results in

$$\frac{P(q^A,p^A,q^B,p^B)}{P(q^B,p^B,q^A,p^A)} = \frac{q^B p^B}{q^A p^A}\,\frac{q^A p^B}{q^B p^A}. \tag{5.21}$$

The multiplication of (5.21) with (4.5) finally yields

$$[P(q^A,p^A,q^B,p^B)]^2 = \frac{q^A p^B}{q^B p^A}\,\frac{q^B p^B}{q^A p^A},$$

i.e., Fisher's "ideal index" (3.3). □

6. Inconsistency Considerations

In Eichhorn/Voeller [1976] we presented the general solution to the so-called inconsistency problem of Fisher's tests. Since these tests form a subset of all the conditions considered in this investigation all previous results concerning inconsistent subsets of conditions are also valid here. But since new conditions have been added new inconsistencies may turn up, too. In the following a very definite viewpoint is taken with respect to these potential inconsistencies. We only look for those inconsistent subsets which result when additional conditions are added to a price index/purchasing power parity of type (2.0). This means, that we assume the fulfillment of Axioms (2.1) to (2.4) (and, hence, according to Remark

(3.10) of Tests (3.11), (3.12), (3.13), (3.14)) by a function P: $\mathbb{R}^{4n}_{++} \rightarrow \mathbb{R}_{++}$ and then ask which further conditions produce an inconsistent set of properties. Therefore we speak of an **inconsistency theorem** or a **non-existence theorem** in connection with an index function of type (2.0) and all other conditions (tests) listed additionally in Sections 3 and 4 if for a subset of these conditions there does not exist a function P of type (2.0) which satisfies all conditions of the subset at the same time.

(6.1) **Remark**:

Inconsistency theorems are at best possible if a condition from $\{(4.10), (4.11), (4.13), (4.14)\}$ and a condition from $\{(4.6), (4.7), (4.9)\}$ are added to Axioms (2.1) to (2.4).

Proof:

Fisher's "ideal index" (3.3) satisfies all conditions except for (4.10), (4.11), (4.13) and (4.14). The Cobb-Douglas Index (3.9) meets all conditions except for (4.6), (4.7) and (4.9). □

Now, the number of possibly inconsistent sets of conditions for an index of type (2.0) is greatly reduced. As a further consequence of the characterizations of Fisher's "ideal index" (3.3) as well as the Cobb-Douglas Index (3.9) it is clear that inconsistencies occur when conditions are added to (2.1) to (2.4) which either index does not fulfill. In particular this implies:

(6.2) **Theorem**:

There does not exist any price index/purchasing power parity of type (2.0) which satisfies either:

(i) the Base Test (4.14) (or the Circular Test (4.13)) and the Determinateness Test (4.9)

or

(ii) the Base Test (4.14) (or the Circular Test (4.13)) and the Product Test (4.7) (or the Factor Reversal Test (4.6))

simultaneously.

Proof:

According to Theorem (5.4), the only function P: $\mathbb{R}^{4n}_{++} \rightarrow \mathbb{R}_{++}$ fulfilling (2.1) to (2.4) and the Base Test (4.14) (or the Circular Test (4.13)) is given by (3.9). This index, however, does not satisfy simultaneously any of the sets of conditions listed under (i) and (ii). □

Of course, any additional condition added to (i) or (ii) also leads to inconsistencies.

(6.3) **Theorem:**

There does not exist any price index/purchasing power parity of type (2.0) which satisfies either:

(i) the Time/Country Reversal Test (4.5), the Factor Reversal Test (4.6), the Price Reversal Test (4.8) and Condition (4.10) (or Condition (4.11))

or

(ii) the Time/Country Reversal Test (4.5), the Factor Reversal Test (4.6), the Price Reversal Test (4.8) and the Base Test (4.14) (or the Circular Test (4.13))

simultaneously.

Proof:

According to Theorem (5.17) Fisher's ''ideal index'' (3.3) can be characterized by Conditions (4.5), (4.6) and (4.8). Since (3.3) does not fulfill the above combinations the theorem holds. □

Footnotes

* We would like to thank Bert Balk and Erwin Diewert for useful comments.

1 Please note the following notation: $x=(x_1,\ldots,x_n) > (y_1,\ldots,y_n)=y$ if $x_1>y_1,\ldots,x_n>y_1$, and $x\geq y$ if $x_1\geq y_1,\ldots,x_n\geq y_n$ but $x\neq y$, and $x\geqq y$ if $x_1\geq y_1,\ldots,x_n\geq y_n$.

2 xy always denotes the inner product of the two vectors x and y.

3 From now on the term ''index'' is used equivalently to ''price index'' and/or ''purchasing power parity''.

4 A quantity index is required to satisfy conditions analogous to those satisfied by a price index/purchasing power parity.

5 As indicated in Footnote 4 the analogous condition for a quantity index would be given by

$$P(p^A,q^B,p^B,q^A) = P(p^B,q^B,p^A,q^A)$$

or, interchanging q^A and q^B, by

$$P(p^A,q^A,p^B,q^B) = P(p^B,q^A,p^A,q^B).$$

References

Diwert, W.E. [1983]. "The Theory of the Cost-of-Living Index and the Measurement of Welfare Change", *Price Level Measurement: Proceedings From a Conference Sponsored by Statistics Canada.*

Drechsler, L. [1973]. "Weighting of Index Numbers in Multilateral International Comparisons", *The Review of Income and Wealth*, pp.17-34.

Eichhorn, W. [1976]. "Fisher's Tests Revisited", *Econometrica* 44, pp.247-256.

________ [1978]. "Functional Equations in Economics", Addison-Wesley Pub. Co., Reading, Mass.

________ and J. Voeller [1976]. "Theory of the Price Index. Fisher's Test Approach and Generalizations", Lecture Notes in Economics and Mathematical Systems. Vol. 140, Springer Verlag, Berlin-Heidelberg-New York.

Fisher, I. [1922]. "The Making of Index Numbers", Houghton-Mifflin, Boston, Third edition revised, 1927. Reprinted by A.M. Kelley, New York, 1968.

Funke, H. and J. Voeller [1979]. "Characterization of Fisher's 'Ideal Index' by Three Reversal Tests", Statistische Hefte, 20. Jg., Hefte 1, pp.54-60.

________ [1980]. "Three Comments on Sato's Paper: The Price Reversal Test and Economic Indices", Statistische Hefte, 21. Jg., Hefte 2, pp.131-132.

________ G. Hacker and J. Voeller [1979]. "Fisher's Circular Test Reconsidered", Schweizerische Zeitschrift für Volkswirtschaft und Statistik, Hefte 4, pp.677-688.

Hacker, G. [1979]. "Theorie der wirtschaftlichen Kennzahl", Dissertation Karlsruhe.

Kravis, I.B. et al. [1975]. "A System of International Comparisons of Gross Product and Purchasing Power", Johns Hopkins University Press, Baltimore and London.

_________ [1978]. "International Comparisons of Real Product and Purchasing Power", Johns Hopkins University Press, Baltimore and London.

Krtscha, M. [1979]. "Über Preisindices und Deren Axiomatische Characterisierung", Dissertation Karlsruhe.

Pollak, R.A. "The Theory of the Cost-of-Living Index", *Price Level Measurement: Proceeding From a Conference Sponsored by Statistics Canada.*

Sato, K. [1980]. "The Price Reversal Test and Economic Indices", Statistische Heft, 21. Jg., Hefte 2, pp.127-132.

Voeller, J. [1982]. "Purchasing Power Parities for International Comparisons", Oelgeschlager, Gunn & Hain, Meisenheim-Cambridge.

COMMENTS

Charles Blackorby
Department of Economics
University of British Columbia

There are two main strands of thought in the study of index numbers. The first is the so-called axiomatic method of constructing index numbers. In this particular approach the researcher considers an index number to be a function of current and base period prices and quantities, and then imposes an axiomatic structure on these index numbers. The result is a narrower class of index numbers than the original set. The other strand of thought approaches index numbers in a similar way except that the quantities, both current and base, are thought to be functions of the prices. Hence, in this alternative, which I'll refer to as the economic theory of index numbers, ultimately the index numbers are functions only of current and base period prices.

In the paper written by Professors Eichhorn and Joachim Voeller we see the high standards which are regularly found in the work of Professors Eichhorn and Voeller, two of the foremost experts in the world of the axiomatic theory of index numbers. In support of the exact economic index number approach, we have papers by Robert Pollak and Erwin Diewert on the Foundations of Economic Index Number Theory. These latter two papers represent the best that the other side has to offer.

Let me be more explicit about the difference between these two approaches. For Professors Eichhorn and Voeller a price index is a function which depends upon four sets of independent variables, that is, current and base period prices and current and base period quantities. This price index is assumed to satisfy four axioms, monotonicity, proportionality, price dimensionality and commensurability. It is important to notice that each of these four arguments is allowed to range independently over the entire positive orthant of Euclidean N-space.

To convert this to an economic index number, à la Diewert and Pollak, assume that the current quantities are the compensated demand functions of a utility maximizing household evaluated at current prices and a specified utility level, and that base period

quantities are the compensated demands of the same household at base period prices and the same utility level. Thus, the index numbers in the exact economic sense are functions only of current and base period prices and a specified utility level (and, if you like, the preference ordering of the household in question).

Let me now return to the paper of Professors Eichhorn and Voeller. This paper is a paradigm of excellence in clarity of presentation, mathematical elegance and scholarship. The results are, however, essentially impossibility results. That is, suppose we restrict our attentions to the class of index numbers mentioned above. Then Eichhorn and Voeller show that if in addition to the above usual regularity assumptions, we require that this price index function satisfy a circularity test and a product test, then the class of possible price indices is empty. That is, there does not exist an index number which satisfies monotonicity, proportionality, price dimensionality, commensurability, circularity and the product test. The results in their paper are correct. In addition Professors Eichhorn and Voeller are very persuasive in the reasonableness of the axioms which they are maintaining. It also provides guidelines for the practical construction of index numbers.

Nevertheless this type of exercise puts me in a rather ''whodunit'' mood. Even though it is the entire set of axioms which is mutually inconsistent, I feel there must be some particular culprit. My candidate is the product test. Not only are prices and quantities independent variables, but the price index and the quantity index are independent variables too. Yet the product test says that they can't be independent.

The two positions outlined above are polar extremes precisely because of the domains of definitions of the index numbers in question. In the case of the mechanistic or axiomatic approach to index numbers, current quantities are independent of all prices and base period quantities are independent of all prices; it is merely the index number itself that must satisfy certain axioms. In the economic index number case current quantities are constrained to be functions of prices which satisfy the Slutsky symmetry conditions by adding up conditions of a single utility maximizing consumer.

What puzzles me about the two approaches, is not that researchers have investigated in great detail these polar cases, but rather that no one has investigated the middle ground. For example, suppose that we begin with Professor Eichhorn's definition of an index number. If we were to relax the assumption that quantities are completely unrelated, independent of prices, we do not need immediately assume that these quantities are the compensated demand functions of a single utility maximizing consumer. It might be sufficient to assume that current quantities are simply the aggregate demand functions of the economy which we are examining. We might then begin to place regularity conditions on these demand functions. The first step would be to assume the demand functions were continuous, homogeneous of degree zero and add to total expenditure. From a general equilibrium system, we can expect no more. We could also assume that the system of aggregate demand equations satisfy the assumption of universal gross substitutes. This, in conjunction with the axioms which Professor Eichhorn finds plausible, would lead to a smaller class of reasonable cost-of-living indices. We have no idea how small such a class might be. However, we do know that the assumption of gross substitutability is sufficient to ensure the uniqueness of competitive equilibrium as well as its global stability. We could go even further without assuming that the quantities in question are generated by those of a utility maximizing consumer. For example, it is known that aggregate demand functions may satisfy some Slutsky symmetry-like conditions but not all of them (under some circumstances); see for example a paper by E. Diewert [1977]. In conjunction with the abovementioned axioms and some Slutsky symmetry-like conditions, the class of feasible index numbers gets larger. Precisely where to stop in this endeavour is by no means obvious, but if we were to systematically study this procedure of making the demand functions satisfy more and more restrictions, thereby enlarging the class of reasonable index numbers, we might learn how to interact with the kinds of demand functions which are actually generated by the real world.

Reference

Diewert, W.E., [1977] " Generalized Slutsky Conditions of Aggregate Consumer Demand", *Journal of Economic Theory*, 15, pp.353-362.

SECTION III

Specific Price Measurement Issues

Problèmes particuliers de la mesure de la variation des prix

ESCALATION MEASURES: WHAT IS THE ANSWER? WHAT IS THE QUESTION?

Jack E. Triplett
Office of Research and Evaluation
United States Bureau of Labor Statistics

SUMMARY

Economists generally believe that a cost-of-living index is the appropriate measure for escalation purposes, a professional judgment that is documented in – to take examples covering a span of two plus decades – the 1961 report of the Price Statistics Review Committee (Stigler Committee) and the 1982 Consultations Feedback Report of the Statistics Canada Price Measurement Review Program. Economists who are specialists in index number theory as well as those who are primarily users of index numbers for research and policy analysis share this opinion.

This paper challenges this view. Escalation, whether in public or private sectors, seldom implies sets of circumstances that correspond to those on which the cost-of-living index has traditionally been defined.

RÉSUMÉ

Les économistes estiment généralement qu'un indice du coût de la vie est la mesure qui convient à l'indexation. C'est un jugement professionnel que l'on trouve formulé, pour prendre des exemples s'étendant sur plus de deux décennies, dans le rapport de l96l du Comité d'examen des statistiques de prix (Comité Stigler) et dans le rapport de 1982 de Statistique Canada sur les consultations relatives au programme d'examen de la mesure

des prix. Les économistes qui sont des spécialistes dans la théorie des nombres-indices ainsi que ceux qui en sont principalement des utilisateurs pour la recherche et l'analyse des politiques partagent ce point de vue.

Le présent document conteste cette opinion. L'indexation, que ce soit dans le secteur public ou privé, implique rarement un ensemble de conditions correspondant à celles d'après lesquelles l'indice du coût de la vie est habituellement défini.

Le chapitre I expose le concept d'indice du coût de la vie, en mettant en relief la multiplicité des mesures qu'entraîne cette notion générale. L'indice du coût de la vie est une réponse à une question économique revêtant la forme générale suivante: "Quelle est la variation minimale d'une variable économique qui serait nécessaire pour qu'une unité particulière de consommation soit indifférente entre un état antérieur et un état postérieur à l'inflation?" On a parfois négligé le fait que les documents publiés portaient sur toute une série d'indices du coût de la vie variant avec la "variable économique" en fonction de laquelle l'indice est défini (par exemple la dépense, le revenu avant impôt ou la richesse). Chacune de ces diverses définitions possibles peut être considérée comme une réponse à une question économiquement pertinente. Par conséquent, il existe de nombreuses réponses prenant la forme d'un indice du coût à la vie à de nombreuses questions relatives au coût de la vie. Chaque question peut être formulée sous l'angle de la compensation de l'inflation; chaque indice du coût de la vie est donc une réponse qui fournit une mesure appropriée à une fin quelconque. Le chapitre I expose ces diverses définitions d'un indice du coût de la vie.

L'indexation des accords salariaux, des pensions alimentaires et des prestations de sécurité sociale, entre autres choses, est souvent considérée aussi comme nécessitant une mesure qui compense l'inflation – interprétation qui est **parfois** *juste. Le chapitre II du document souligne cependant que, même lorsque l'indexation a pour objectif de compenser l'inflation, l'enjeu réel correspond rarement à la question de compensation qui est inhérente à toute formulation classique d'un indice du coût de la vie. C'est peut-être une des raisons pour lesquelles les parties aux ententes d'indexation semblent peu intéressées par la conception que les économistes se font d'un indice du coût de la vie. Le document présente une mesure des prix* **à la façon** *du coût de la vie qui convient aux fins de l'indexation et évoque les problèmes posés par la formulation et l'estimation d'un "indice d'indexation".*

Le dernier chapitre relève que l'objectif recherché dans certains cas importants d'indexation, comme celui des prestations de sécurité sociale aux États-Unis, n'a pas été bien spécifié. Autrement dit, quelle est la question précise à laquelle on cherche à répondre? Comme on s'est servi de mesures de prix pour l'indexation, on a pris pour acquis que l'objectif de ces dispositions d'indexation était la protection contre l'inflation; or, le débat public (on est tenté de dire "l'insatisfaction publique") au sujet des résultats de l'indexation porte à croire que l'enjeu est dans une large mesure la détermination de l'équité au niveau des revenus réels. Cela implique d'autres questions économiques et des mesures économiques différentes du concept d'indice du coût de la vie sur lequel les débats ont tellement porté jusqu'ici. Le dernier chapitre traite de la formulation des objectifs d'indexation et de la spécification des mesures – principalement des salaires et des revenus – qui répondraient à des objectifs de remplacement dans les dispositions d'indexation.

Introduction

Economists generally believe that a cost-of-living index is the appropriate measure for escalation purposes, a professional judgment that is documented in – to take examples covering a span of two plus decades – the 1961 report of the Price Statistics Review Committee (Stigler Committee) and the 1982 *Consultations Feedback Report* of the Statistics Canada Price Measurement Review Program. Economists who are specialists in index number theory as well as those who are primarily users of index numbers for research and policy analysis share this opinion.

This paper challenges this view. Escalation, whether in public or private sectors, seldom implies sets of circumstances that correspond to those on which the cost-of-living index has traditionally been defined.

Section I reviews the cost-of-living index concept, emphasizing the multiplicity of measures that the general concept implies. The cost-of-living index is an answer to an economic question, of the general form: "What is the minimum change in an economic variable that would be required in order to leave a specified individual consuming unit indifferent between pre- and post-inflationary states?" It has sometimes not been recognized that the

literature encompasses a whole family of cost-of-living indexes, which vary with the ''economic variable'' on which the index is defined (e.g., ''expenditure'', ''pre-tax income'', ''wealth''). Each of these alternative definitions can be thought of as the answer to an economically meaningful question. Thus, there are many cost-of-living index answers because there are many cost-of-living questions. Each question can be thought of in terms of a compensation for inflation, and each cost-of-living index is an answer that provides an appropriate measurement for some purpose. Section I of the paper discusses these alternative definitions of a cost-of-living index.

Escalation of collective bargaining agreements, divorce settlements, social security payments, and so forth is also frequently interpreted as requiring a measure that will compensate for inflation, and this interpretation is **sometimes** correct. Section II of the paper points out, however, that even when the escalation objective is inflation compensation, the question implied by escalation seldom corresponds to the compensation question that is inherent in the traditional cost-of-living index formulations. That is perhaps one reason why participants in escalation arrangements seem uninterested in the economists' concept of the cost-of-living index. The section develops one cost-of-living **like** price measure that is relevant to escalation, and discusses the problems of formulating and estimating an ''escalation index''.

The final section of the paper notes that the objective sought in some prominent escalation situations, such as the escalation of social security payments in the U.S., has not been specified clearly. To put it another way, what is the precise question for which an answer is being sought? Because price measures have been employed for escalation, the presumption has been created that the objective of these escalation arrangements is inflation-protection; yet, much of the public discussion of (one is tempted to say ''public dissatisfaction with'') the outcome of escalation suggests issues involving the determination of equity in real incomes. This implies alternative economic questions and alternative economic measurements to the cost-of-living index concept on which so much of the discussion has focused in the recent past. The paper concludes by discussing the formulation of escalation objectives and the specification of measures – primarily of wages and incomes – that would meet alternative objectives in escalation arrangements.

I. Cost-of-Living Questions and Cost-of-Living Indexes

The development of the cost-of-living index concept occurred largely because of a need to make precise the questions that a price index was to answer. Starting with the original formulation by Konüs, the idea has gradually spread to become a standard part of the economists' intellectual tool kit, and a favorite artifact of textbooks in microeconomic theory.

The very universality of the cost-of-living (COL) concept, however, cloaks the fact that economists use the term "cost-of-living index" in slightly different ways. Until the last decade or so, most of the literature was written as if there were but one theoretical COL index and that it provided the "true" index against which approximations such as the Consumer Price Index (CPI) were to be assessed. On this view, if the usual textbook presentation – usually cast in an indifference curve-budget constraint diagram – was conceded to be oversimplified, removing the oversimplifications was still presumed to leave one generally accepted COL index definition.

The COL index concept is usefully thought of as framing an economic question (to which the COL index itself is the answer). That question is usually phrased in a manner similar to the following (the language is adapted from Samuelson and Swamy, [1974] p.567): "What is the ratio of the (minimum) costs of a given level of living in two price situations?"

Once attention is drawn to what one might call the "COL question", it becomes apparent that a family of precisely stated questions exists, and not just a single question. As there are many appropriate and relevant COL questions, there are in consequence many formulations of a COL index, not just a single one. In a way this is not surprising, for one of the oldest adages of index numbers holds that the design of the index depends on its purpose. One can distinguish at least the following family members.

Expenditure-defined COL index. The most frequently used (and most straightforward) COL index is the formulation that answers the question: "What ratio of expenditures is required to maintain a fixed living standard in two price regimes?" This formulation may be termed the "expenditure-defined" COL index.

It is well known that the question underlying the expenditure-defined COL index can be framed from a variety of **perspectives**, depending on the particular living standard that is chosen. What one might refer to as a ''Laspeyres-perspective'' expenditure-defined COL index holds constant the base, or ''reference'', period living standard – that is, following Pollak [1971] the reference period preference function and the indifference curve attained in that period. The expenditure ratio then gives the change in cost that would maintain the consuming unit at the reference period living standard.

An alternative is the ''Paasche-perspective'' expenditure-defined COL index that is based on the current, or ''comparison'', period living standard. Pollak [1971] points out that other living standards may also be relevant; e.g., the change in cost between the years 1979 and 1982 of the living standard attained in 1972, or the comparison of the cost of a U.S. living standard in Norway and Egypt. Because both intertemporal and interarea COL indexes are normally wanted for three or more periods or places, and not just for the two period comparisons encountered in textbooks, this point has greater significance for price index measurement than is sometimes recognized.

Much of the content of traditional price index theory concerns the effects on the measurement of what the previous paragraphs have termed the ''perspective''. As these topics are adequately treated elsewhere in this conference, there is no need to explore them here. Two points, only, must be made, not because either is new but only because they seem so widely misunderstood.

First, various alternative perspectives all give equally valid measures. They deal with subtly different questions, and apply to subtly different uses. If one wants to compare (say) 1972 and 1982 prices, and if it were true (as often alleged) that the market basket for 1982 is far different from that of 1972, or that 1972 and 1982 correspond to different living standards, this would not automatically mean that the 1982 perspective was always preferred to the 1972 perspective – it depends on the uses and the questions asked. This point deserves emphasis because it has so frequently been misunderstood.[1]

Second, it is natural to take the ''two price regimes'' mentioned in the cost-of-living question as consisting of the relevant market prices prevailing in the two periods being

compared, and not prices from some other periods. That is, a COL index for the current month (or current year) should depend on market prices prevailing **this month** (or this year), plus those of the reference period, and not those contracted for at some other period. One often hears contrary statements, such as: "The COL index for persons who own their own homes (or refrigerators or cars) should not incorporate current house prices (or refrigerator prices or car prices), because current price movements do not affect that person's payments for durable goods acquired in some previous period" (for an example, see Kahn, 1980). Such statements reflect failure to recognize that it is the consumption of the services of durable goods that matters in a COL index, and not the acquisition of the durable goods themselves (a distinction that has prevailed in the analysis of consumer demand at least since the publication of Harberger, 1960), combined with failure to understand that current opportunity cost, and not historical replacement cost, defines the consumer's opportunity set (though there are, to be sure, ambiguities for the measurement of **one period** cost levels that are discussed in Muth, 1974 and Pollak, 1975a).[2]

The expenditure-defined COL index is the workhorse model for empirical estimates. The U.S. Consumer Price Index is an approximation to the expenditure-defined COL index.[3] The expenditure-defined COL index is also the model that has been employed for all empirical COL indexes that have been based on estimated sets of consumer demand relations (see, for example, Braithwait, 1980; Christensen and Manser, 1976; Manser, 1975; Goldberger and Gameletsos, 1970).

The expenditure-defined COL model was a natural choice for empirical COL index estimation, since nearly all research in consumer demand systems has used total expenditures, rather than "income", or some other variable as the consumer's budget constraint (see the survey by Brown and Deaton, 1972, or Phlips, 1974, or Deaton and Muellbauer, 1980). In part, this research strategy reflects a decision to avoid intertemporal decisions inherent when saving is admitted into the consumer's decision-making problem, complications that are usually deemed tangential when the research focus is on the allocation among goods for current consumption; and in part it reflects the reality that available survey data on expenditures are usually considered more reliable than those on saving and income. These same considerations – avoidance or minimization of certain technical problems, and the desire to erect the estimates on the firmer part of available consumption data – lie behind the decision to employ the expenditure-based COL index model for the CPI.

This "defense" of the expenditure-defined COL index is necessary because it is plainly the most limiting of the family of COL indexes. That both economic researchers and statistical agencies have stuck with the expenditure-defined concept despite the attractiveness of the alternatives says a good deal about its homely virtues.

Income-defined COL index. This COL index gives the answer to the question: "What ratio of (pre-tax)[4] incomes would be required to maintain a fixed standard of living in two price regimes?" As with the expenditure-defined COL, there are alternative perspectives corresponding to the living standard that is used for the comparison (living standards from either reference period or comparison are those usually considered), and different perspectives may produce different measures. The income-defined COL also uses as data market prices for the two periods under consideration, as does the expenditure-defined COL.

An income-defined COL index would differ from an expenditure-defined COL index in a number of ways. One important difference is in the treatment of income and payroll taxes. In an expenditure-defined COL index (and in the CPI, which is its approximation) an increase in these taxes does not affect the index, though a change in excise or sales taxes does. This gives rise to anomalies if one form of tax is substituted for the other. The income-defined COL index has the advantage that it would not be affected by the mix of income and sales taxes, but only by their combined amount.

An income-defined COL index has been considered preferable to an expenditure-defined index for many purposes. It has been argued, for example (see Cagan and Moore, 1981) that an income-defined COL is more appropriate for escalation use, on the grounds that the entity being escalated is an income source, and not a measure of consumer expenditure (in Section II, we examine these grounds).

However, the CPIs of most countries approximate an expenditure-defined COL index and not an income-defined one. And even research estimates of an income-based COL index concept are infrequent (see the Gillingham and Greenlees paper presented at this conference). Two attributes of the income-defined COL account for its rarity.

First, the income-based COL would rise with an increase in (say) individual Social Security

payroll taxes, even if there were no change in any price in the economy. For some purposes, this would prove objectionable. Second, the concept of income, though seemingly simple at first glance, is notoriously difficult to define and measure in economic terms, and as a concept merges into lifetime wealth without a clear demarkation;[5] thus the greater **apparent** usefulness of the income-based COL is offset by formidable measurement problems in practice. Additional discussion of the income-defined COL index is in Gillingham and Greenlees [1983].

The Non-market Commodities COL index. Implicitly, the COL indexes discussed previously were defined on goods and services acquired through the market (or that could be acquired through markets – as for example, imputation in the U.S. CPI of the value of homegrown food, and beginning in 1983, of the rental value of owner-occupied housing). The standard of living may also depend on the level of services provided by the government, and on aspects of living such as pollution. Once this distinction is recognized, then the numerator of the expenditure-defined COL index discussed above should be rephrased to answer the question: ''What is the cost, at today's market prices, of a bundle of **market-purchased** goods and services equivalent to the bundle consumed in the base period?'' Analogous rephrasing can also be made for the income-defined COL index.

The ''non-market commodities COL index'' is a more comprehensive COL concept, dealing with the question: ''What change in cost (alternatively, what change in income) is required to maintain the base period's living standard, considering privately-provided goods and services, and also free government services, the effects of pollution, and so forth?'' As the alternative wording of the COL question makes clear, a complete taxonomy of the family of COL indexes would include both expenditure-defined and income-defined variants of the non-market commodities COL index, and the varying perspectives that were discussed in the previous sections also arise on this definition.

Since the non-market commodities COL index encompasses a more comprehensive set of consumption commodities than was the case for the regular expenditure-defined and income-defined COL indexes discussed earlier, each of the latter is a ''sub-index'' (in the sense of Pollak, 1975b) of the relevant version of the non-market commodities COL index. The theoretical discussion of sub-indexes (see also Blackorby and Russell, 1978) can

be expected to apply to the relationship between the sub-indexes for market-purchased commodities and the more comprehensive index.

There can be no doubt of the relevance and usefulness of the non-market commodities COL index. The empirical barriers to estimating it, however, are formidable, because it requires consumer valuations of non-market commodities (the ''value of clean air'').

Some research on this topic has been carried out. Several years ago the BLS explored the possibilities in the ''median-voter'' literature and some alternative approaches, but we concluded that information that could be extracted from these approaches did not satisfy the requirements of a COL index (see Cobb, Barkume and Shapiro, 1978, and Shapiro and Smith, 1981). I should note also that the extensive literature using hedonic methods to estimate the ''demand'' for neighborhood amenities (including pollution) is defective for our purposes (indeed, for most purposes), for the methodological reasons outlined in Brown [1983] and Triplett [1983a]. So far, empirical estimation of any form of the non-market commodities COL index has proven intractable.

The narrower COL indexes (as, for example, the expenditure-defined COL index) may pick up some of the consumer costs of changes in non-market commodities. For example, if an increase in air pollution causes an increase in medical expenditures because of respiratory illness, part of the consequences of air pollution would show up in the normal expenditure-defined COL index. But for this measure to provide a correct estimate of the value of the non-market commodities COL index would be fortuitous and unlikely. Similarly, it has sometimes been argued that the cost of putting smog control devices on automobiles (which has clearly increased the price of cars and the cost of automobile transportation, and therefore the expenditure-defined COL index) ought to be adjusted out of a COL index because the value of cleaner air provides an offset to the increased private cost of transportation. Note that this suggestion is incorrect if the expenditure-defined COL index is the subject of discussion; and if the non-market commodities COL index is the one that is wanted, then the proposal will approximate the correct movement in that index only if smog regulations are chosen so that the marginal cost of smog abatement equals the incremental valuation on clean air. That this is an appropriate principle for regulation does not mean that it has been met in practice.

The preceding was somewhat of a digression from the main line of the argument, and was intended to illustrate the formidable information requirements of the non-market commodities COL index, and clarify the relations among the various COL index concepts (on which there has been much confusion). At present, I know of no empirical estimate of a COL index including non-market commodities.

Wealth-defined COL index. A COL index based on a wealth measure has been suggested by Alchian and Klein [1973], the objective being to bring assets and changes in asset values into the COL analysis. This idea seems appealing; for one thing, permanent income is a wealth concept, and other income concepts prove slippery or not economically relevant. Explicitly defining the COL index measurement on wealth is a way of cutting through to the essentials. In addition, the wealth concept gets away from the one period decision-making mode that underlies other COL indexes.

An alternative approach to introducing intertemporal decision-making into the theory of the COL index is Pollak [1975a]. Pollak's approach shows that, far from making the measurement easier, moving to a multi-period setting makes it far more difficult. Discussion of these problems in the present paper takes us too far afield.

Partial-income COL index. So far, each successive member of the family of COL indexes has widened the variable on which the index is defined. One can also usefully consider going the other way. This approach is inspired by Pencavel [1977], who sought to determine the minimum change in a **single price** in the consumer demand system that would be sufficient to compensate for the net effect of changes in the other prices. Pencavel's objective was to make the analysis apply to an individual who was a consumer of all commodities and a seller of one.

For present purposes, we can alter this approach a bit by supposing only that an individual has two (or more) sources of income (IT = IA + IB). Then we may ask the question: "What change in IA is required to achieve a **total income** (IT) that maintains the base period living standard?" The resulting "partial-income COL index", being the escalation of IA required to hold utility constant, obviously depends on what has happened to IB.

This measure would probably not normally be properly thought of as a COL index at all, but it does provide a form of escalator, which is one use often proposed for a COL index. The partial-income COL index is the appropriate escalator for situations in which only a portion of income is escalated and the purpose of the escalation is to maintain living standards. We return to this measurement concept in Section II.

Summary. This taxonomy of COL indexes is doubtless not complete, but that is not the intention. The variety of COL indexes that can be produced (expenditure-defined, income-defined, non-market commodities, wealth-defined and partial-income were the names given to the five concepts discussed) correspond to different versions of the question for which the COL index is conceived as the answer. Though some of these questions are more interesting and meaningful than others, there is generally a trade-off between comprehensiveness and practicality in making a COL index measurement. The following sections consider the use of the COL index concepts discussed in this section as escalation measures.

Comment. It is often said that the CPI ''is not a COL index'' because the latter would account for factors such as income taxes, pollution and government services and so forth that are omitted from the CPI (see, for one such statement, Cagan and Moore, 1981, p.1). The discussion in this section specifies the precise sense that such statements are true: they are correct if the broader definitions of a COL index (the income-defined COL index or the non-market commodities COL index) are meant. However, the CPI is not deficient in these elements with respect to the expenditure-defined COL index, which provides its theoretical underpinnings. A more precise and less confusing way of putting the matter would be to say that an expenditure-defined COL index omits factors such as income taxes, government services and so forth, whose inclusion in a COL index would be useful for many purposes. Because I believe the expenditure-defined COL index is also relevant and useful for some purposes (especially as a design for an analytic inflation measure – but that is beyond the scope of the present paper), one should avoid confusing the choice of the COL question that one wants to answer (which implies the definition of the COL index that one wants to compute) with the issue of how well the CPI approximates its own COL index concept. Both are important matters; but they are distinctly different ones.

II. Escalation or "Indexing" Issues and Escalation Measures

In 1970, benefits of programs that accounted for roughly 3 percent of U.S. Federal government outlays were tied to the U.S. Consumer Price Index. During the following decade more and more government programs were "indexed", and by 1980 this proportion had risen to 30 percent. It should be noted that this great percentage increase reflects not so much growth in programs that are indexed (though that has occurred) but rather growth in the number of programs that make use of indexing (see Appendix A of DeMilner, 1981, or Table 4 of Goldfeld-Ooms, 1981). In Canada payments made under a comparable group of social programs (including pensions, family allowances, among others) are adjusted by the Canadian Consumer Price Index, which is also used to index income tax schedules (see Statistics Canada, 1982, p.77).

Private sector use of the Consumer Price Index as an escalator in collective bargaining agreements has fluctuated with the rate of inflation in the post-war period; in the 1970s, the number of U.S. workers covered by escalator clauses expanded as the rate of inflation increased. Few private sector collective bargaining agreements provide for 100 percent escalation, and the data available suggest that the escalator yield as a proportion of the CPI is an inverse function of the CPI's rate of change (see Cagan and Moore, 1981, Table 2). Canadian collective bargaining contracts make use of the Canadian CPI in similar ways. One should note that at least one collective bargaining agreement (that of the United Auto Workers) uses an average of the U.S. and Canadian CPIs.

Outside the traditional collective bargaining use, escalation in private sector agreements has also grown greatly in recent years. Though I know of no concrete data to indicate the degree of usage, fragmentary information in the BLS (frequently, a letter of inquiry occasioned by a dispute in interpreting an often unclearly written agreement) indicates that divorce settlements, rental agreements, and so forth have increasingly been tied to the Consumer Price Index. Comparable private-sector uses of the CPI are reported for Canada (Statistics Canada, 1982, p.77). A novel and perplexing use of the CPI was the U.S. Financial Accounting Board's decision that the index should be used for deflating data in corporate financial statements in order to get a profit measure that was not distorted by inflation.

What is the purpose of escalation or indexing income payments in private contracts and government transfer programs? What is the appropriate measure for use in such arrangements?

In response to the first question, most people would answer that the objective was to protect workers and benefit recipients from inflation. And given the answer to the first question, economists invariably respond that the cost-of-living index is the appropriate measure for that purpose.

Considering particular cases of escalation leads me to challenge that ''inflation protection'' answer to the first question. The answers that people give to questions about the motivation of their economic behavior characteristically suggest behavior different from what is actually observed. I believe the ''inflation protection'' answer is, though not necessarily wrong, quite incomplete.

But let me put aside the first question until Section III, and consider for the moment only the second question (''What is the appropriate escalation measure?''), as if the answer to the first question were indeed only ''inflation protection''. ''Protection'' against inflation implies compensation. The COL question can be thought of in similar terms – i.e., What change in some economic variable is necessary to compensate for inflation? The issue, then, is whether the variable that is being escalated corresponds to the variable on which some member of the family of COL indexes is defined.

Of the list of COL indexes described in Section I, it is immediately evident that most are defined on variables that differ from the ones chosen for actual escalation situations. We have already noted Cagan and Moore's contention that the expenditure-defined COL index (or its CPI approximation) is inappropriate for escalation because no known escalator is applied to consumer expenditures: ''It [the index] covers only consumption expenditures and not the part of income that is taxed or saved, whereas escalation is directed to income without regard to its disposition'' (Cagan and Moore, 1981, p.1).

But do we in many actual situations escalate income? There may be a few cases in which this is done, but the most prominent escalation cases cannot be so interpreted. For exam-

ple, the average (or ''representative'', to follow the conventions of price index theory) Social Security recipient has some form of income other than Social Security payments. That means we are escalating a portion of income, and not all of it. And perhaps some workers who are covered by cost-of-living escalator clauses under collective bargaining agreements have only income from wage and salary earnings under those agreements. But not all do, and a very large proportion of workers do not fall in the single-source-of-income class if we consider how conventional it is to view owner-occupied housing as producing an imputed income in kind. In all these cases total income is not the variable being escalated, and therefore the income-defined COL index is not the escalator that will leave the individual exactly compensated for inflation. That is, if the income-defined COL were applied as an escalator to a particular payment stream, the resulting total income would not in general equal the income that would keep the individual consuming unit on a fixed indifference curve.

When the particular stream of payments being escalated is only a portion of income, and the objective of escalation is to protect the real living standard, then the partial-income COL index is the appropriate escalator. This index would escalate (say) Social Security benefits by an amount just sufficient to maintain the total real income of the average recipient, after accounting for changes in the recipient's other sources of income (including imputed income from owner-occupied housing).[6]

Yet, there is something unsatisfactory about proposing the partial-income COL index as an escalator. It seems doubtful if Congress, employers, workers or Social Security recipients would find satisfactory an escalator that made compensation for services under collective bargaining contracts, or the level of Social Security benefits, depend on what happens to other income sources – as does the partial-income COL index. Though the partial-income COL index gives the precise answer to the relevant compensation question, perhaps this degree of precision is not what was wanted or is not understood. As Samuelson and Swamy (1974, p.587) remarked in a similar context: ''Probably, though, one should not try to read anything so definite into people's vague notions of equity.''

Moreover, it is not always certain that ''maintaining living standards'' is precisely the meaning of ''inflation protection''. One frequently hears statements such as: ''The pur-

pose of escalating benefits is to maintain the purchasing power of benefit payments, or to maintain the standard of living'', where it is clear from the context that the speaker is under the presumption that these are alternative expressions for the same thing. Where full income is being escalated, they of course are. But where the benefit payment is only a portion of total income, ''maintaining the purchasing power of benefits'' is not the same thing as ''maintaining the standard of living'' – they are different objectives, they imply different escalators, and a choice must be made between them.

Where benefits are only a portion of income, ''maintaining the standard of living'' is a precisely-defined objective that implies the use of the partial-income COL index for escalation. ''Maintaining the purchasing power of benefit payments'', on the other hand, presumably implies that the benefits being escalated should command a constant level of real goods and services – presumably consumption, but sometimes this is not entirely clear – irrespective of the command over goods and services that accrues to total income. For ''maintaining purchasing power'', the expenditure-defined COL index, its fixed-weight approximation, or some other measure, may be suitable. It is evident that the partial-income COL index, reflecting as it does the movement of other income sources, could hardly be interpreted as ''maintaining the purchasing power of **benefits**''.

In summary, the correct escalator for benefits depends on the objective being served by escalation. The COL index concept, seemingly so precise and theoretically appropriate, does not in any of its usual forms match the variables that are known to be escalated in actual situations. The partial-income COL index has properties (mainly, its dependence on other income sources) that make it doubtful that this concept is generally what is wanted. That brings us back to the question that led off this section: What is the purpose of escalation? We turn to this matter in Section III.

III. Matching the Design of an Escalator to the Purpose of Escalation

We now consider the logically prior question: ''What is the purpose of escalation?'' I freely concede that I do not know the answer to this question with any degree of certainty.

This section does not so much propose an answer or answers, but seeks to highlight the question. Much of the recent search for "answers" to the problems posed by "indexing" (of Social Security payments, for example) has taken place without sufficient attention to the questions for which answers are being sought.

It is clear that whatever the purpose people had in mind in entering into escalation arrangements, much dissatisfaction with the results developed during the peak inflation years of 1979-1981. Escalation and indexing arrangements were judged to cost too much, and indeed payments under escalators did rise steeply and in many cases unexpectedly. Not surprisingly, the situation led many to challenge the validity of the price measure (usually the CPI) used as an escalator – see for example, U.S. Congress, **Hearings** [1980] or Statistics Canada [1982]. The accuracy of the CPI as an inflation measure, or as an approximation to a cost-of-living index, is of course a legitimate question, but one that will not be considered in the present paper, because the issues have been adequately covered elsewhere (in Triplett 1983b I reviewed three studies of the U.S. CPI that were commissioned at the height of concern over indexing – Cagan and Moore, 1981; DeMilner, 1981; and Goldfeld-Ooms, 1981).

But even though much of the dissatisfaction with the outcome of indexing has been directed toward the escalating index, many of the issues raised suggest that the speaker has in mind an objective that does not match the purposes of the CPI or of a cost-of-living index. Examining carefully some of these complaints can tell us a good deal about what is desired in the situations for which indexing has been employed, and thereby lead to better specifications for measurement.

Lowering Contracting Costs. It has become a commonplace observation that when Congress tied U.S. Social Security benefits to the CPI, its objective was to **lower** the rate of increase in per person benefits. The evidence suggests that it succeeded. Goldfeld-Ooms [1981] present data indicating that the increase in Social Security benefits per recipient rose more rapidly than the CPI in the decade before "indexing" was adopted in 1975; benefits per recipient have risen more slowly than the CPI since that time, so that real benefits per recipient dropped 5 percent between 1975 and 1981.[7] A major objective behind the

decision to use escalation in Social Security benefits was to get an emotional and politically explosive issue out of the legislative arena.

A similar point can be made about collective bargaining agreements. Are employers and unions attempting only to assure that workers' real incomes are precisely protected from inflation? Or is the foremost objective to remove from the collective bargaining table an emotionally-loaded and difficult and expensive-to-negotiate issue? Reducing the cost of bargaining may be as strong a motivation as the explicit and exact protection of the workers' standards of living.[8] This in turn suggests that because there are multiple objectives to be met in an escalator clause, precise definition of the cost-of-living index to be used in the contract becomes of second order importance.

Of course, one might object that the workers would only agree to remove the inflation issue from the bargaining table if they were fully protected; but even if this were true of workers, an escalator clause puts management in the position of speculating on the course of the CPI. Putting an upper limit on losses from such speculation is one reason why collective bargaining agreements typically do not provide 100 percent pass-through of the CPI rate of change. It may also be one reason why no party to a contract agreement, so far as I can determine, has ever specified that the economist's concept of a COL index is wanted, even when the distinction between a fixed-weight price index and a COL index has been described to them – and the workers, judging from positions taken by their representatives, strongly dissent from the idea that any COL index is relevant to collective bargaining (see Oswald, 1980). Unless contracting parties are not acting in their own interests (which I doubt), we should probably pay more attention to what motivates the adoption of an escalator clause in determining the statistical formulation that is appropriate.[9]

Income-Equity Issues. One characteristic of recent criticism of indexing is its focus on what has happened to the incomes of the population that receives index payments. As Cagan and Moore [1981], put it, the indexed population is insulated from ''price changes that reflect a change in the standard of living of the entire population'' (p.3). Thus, if a price increase in imported oil leads to a fall in national income, the larger is the protected, indexed part of the population, the larger the decline in real income that must be taken by the non-indexed population. Of course, if the entire population were indexed, the situation would be disastrous.[10]

Reaction to what many thought was "petro inflation" in the U.S. and Canada has resulted in suggestions for a measure of "domestic inflation" as a remedy for what many people regard as the overadjustment of recipient incomes from use of presently available price measures (see Statistics Canada 1982). Cagan and Moore [1981] propose (p.3) that the CPI be adjusted by the ratio of U.S. export and import prices, so long as the rate of price change exceeds the rate wage change.

I have reservations about whether one can really distinguish "imported" and "domestic" inflation. Aside from this, however, is a more basic reservation: the problem complained of concerns equity in the distribution of real income. If distributional equity is the goal that is desired, the solution is to assure that the incomes of the recipient population change with the incomes of the remainder of the population (e.g., for Social Security beneficiaries, to tie benefits to a measure of wages). Trying to pursue equity in the distribution of real income by devising some adjustment to the deflator for income (the CPI) is not only an exceedingly cumbersome approach, it confuses the objective of equitable growth in real income with that of inflation protection (which means constant real income). Equitable sharing of income declines and maintaining income constant are different objectives that cannot be reconciled by any indexing arrangements.[11]

On the other hand, linking growth in individual retired incomes to the growth in per capita incomes of workers may well be unfeasible if the proportion of beneficiaries in the population rises. This is a very real threat to the U.S. retirement system because of the population bulge of 30-40 year-old workers now moving its way through the demographic structure. This suggests an alternative objective for escalation plans: that the share of national income going to the total beneficiary population be limited at some level, lest the tax burden on the working population reduces incentives to the point that it affects productivity growth. If the rationale underlying a Social Security system that pays out to the average beneficiary more than the actuarial value of his payments into the fund is that such a system shares current productivity gains with the retired population, then one cannot tolerate a beneficiary scheme that threatens these productivity gains.

Either way of looking at the problem suggests alternatives to current indexing methods – an indexing proposal that constrains payments to beneficiaries so that the share of na-

tional income to Social Security recipients be held below some ceiling, perhaps by a formula that sets payments to each beneficiary on the basis of wages, adjusted by the share of the per capita productivity dividend that is to be allocated to the total retired population (and the latter could produce a negative adjustment for individual retiree payments if the growth rate of the retired population exceeds the growth rate of productivity).

These are intended only as examples to illustrate the proposition that thinking through the goals to be met by a situation in which indexing is proposed leads to a more precise definition of the question that indexing is to answer. Once the question has been refined, the development of a measure that would answer the question is a technical task. To be sure, the formulation of indexing objectives is by far the most difficult task, and it is not entirely an economic one. Much of the recent debate over indexing seems preoccupied with finding some narrow technical measure that would obviate the necessity for making the difficult choices that society must make. This is, of course, impossible. And the very search for a narrow technical ''answer'' is counterproductive and postpones work on the real task: What is the question?

Footnotes

[1] Fisher and Shell [1972] argued that **in the presence of taste change between two periods** using the Paasche-perspective index can permit framing a meaningful question (whereas economists have conventionally disclaimed the possibility of making comparisons that bridge taste change, the COL index being defined on an unchanging preference map). The text does not relate to the Fisher-Shell position.

[2] One could, it is true, devise a measure of "current outlays" compared with reference period outlays, for cases where consumer-unit payments differ from market prices prevailing in that period. One attempt was the experimental outlay measures labelled "X-4" and "X-5" published in the U.S. Consumer Price Index press releases beginning in 1980. Framing the precise question for which such measures provided answers, however, does not yield a straightforward result, suggesting that the reasons for proposing such measurements have not been thought through carefully.

[3] See Gillingham [1974] for a discussion of the ways that the theory of the cost-of-living index has guided decision-making in the construction of the U.S. CPI.

[4] One might distinguish between pre-tax income and after-tax income in computing this member of the COL index family – that is, a complete taxonomy of the COL index family tree would do so – but the more interesting case is the index involving pre-tax income.

[5] Insight into the difficulties of defining a meaningful concept of income and its measurement is contained in Friedman's Theory of the Consumption Function.

[6] The way consumption is defined has implications for the variable on which escalation takes place. Consider, for example, the treatment of durable goods in a COL index. It is very clear that a consumption measure requires data on the services of durable goods such as owner-occupied housing. To use for escalation a CPI incorporating a flow-of-services treatment of housing implies that the imputed income from owner-occupied housing be incorporated into the income measure being escalated. If it is not, then the partial-income COL index becomes the relevant escalator, since it explicitly considers the movement of income sources other than the one being escalated.

[7] This estimate is based on my own updating of Goldfeld-Ooms Table 2 (which shows a 3 percent fall in real benefits between 1975 and 1980). Both estimates use the official CPI, so the decline would be less if real benefits were computed by use of a price index that treated housing from a flow-of-services approach (such as Gillingham, 1983).

[8] In this regard, I have observed that groups of union members generally think that their own inflationary experience exceeds that measured by the CPI. National opinion polls suggest that a large part of the public in the U.S. holds a similar belief, at least for recent periods. Thus, when inflation "catch-up" or inflation expectations become a factor at the bargaining table, differing perceptions of the facts can extend negotiations.

[9] During the past several years, statistical agencies have often been criticized for not providing guidelines for writing escalator clauses, or for their failure to tell the parties how escalator clauses should be written or what measurements should be used. Much of this criticism stemmed from the belief that the CPI was an independent contributor to inflationary pressure (an "engine of inflation", Alfred Kahn called it) and that changing the index or the way it was used could somehow moderate inflation in the private sector. This criticism is misguided. Parties to a collection bargaining agreement will determine what escalator provision they believe is appropriate. No govern-

ment agency, and certainly not a statistical agency, has a role in this decision. Moreover, it is naïve to believe that changing the numbers used in an escalator clause formula will have any but the most transitory and ephemeral effects on the course of money wages or inflation.

10 It seems such a short time ago that many economists argued that the cost of inflation could be mitigated, provided all portions of the population were indexed, and some of them pointed, paradoxically, to hyper-inflations such as Brazil and Israel as examples of how indexing might work!

11 A frequently heard objection to the use of a wage measure for indexing purposes is that a wage measure would share gains in productivity with the retired population, which, it is alleged, does not contribute to productivity gain. This objection seems predicated on a false premise. The current level of productivity depends on the contributions of past generations, and some part of the current capital stock, both physical and human capital, was accumulated by current retirees. Unless retirees have already appropriated the returns from all the social investments they may have undertaken in their working lifetime, a system, such as the U.S. Social Security system, that makes the pension depend on the current social dividend as well as the pensioner's own contribution to the account has much to recommend it on equity grounds. The precise division of National Income among working and retired population is a political issue.

References

Alchian, Armen A. and Benjamin Klein [1971]. "On a Correct Measure of Inflation". *Journal of Money, Credit and Banking* 5:173-91.

Blackorby, Charles and R. Robert Russell [1978]. "Indices and Subindices of the Cost-of-Living and the Standard of Living". *International Economic Review,* Vol. 19 (February), pp.229-40.

Braithwait, Steven D. [1980]. "The Substitution Bias of the Laspeyres Price Index: An Analysis Using Estimated Cost-of-Living Indexes". *American Economic Review* 70:64-77.

Brown, Alan and Angus Deaton [1972]. "Surveys in Applied Economics: Models of Consumer Behavior". *Economic Journal* 82:1145-1236.

Brown, James [1983]. "Structural Identification ahd Hedonic Markets". In Jack E. Triplett, ed., *The Measurement of Labor Cost.* Conference on Research in Income and Wealth, Studies in Income and Wealth, Vol. 48, Chicago, Illinois: The University of Chicago Press for the National Bureau of Economic Research.

Cagan, Phillip and Geoffrey H. Moore [1981]. *The Consumer Price Index: Issues and Alternatives.* Washington, DC: American Enterprise Institute for Public Policy Research.

Christensen, Laurits R. and Marilyn E. Manser [1976]. "Cost-of-Living Indexes and Price Indexes for U.S. Meat and Produce, 1947-1971". In Nestor E. Terleckyj, ed., *Household Production and Consumption.* Conference on Research in Income and Wealth: Studies in Income and Wealth, Vol. 40. New York: National Bureau of Economic Research.

Cobb, Steven, Anthony Barkume and Perry Shapiro [1978]. "Amenities and Property Values in a Model of an Urban Area: A Comment". *Journal of Public Economics* 9:107-110.

Deaton, Angus and John Muellbauer [1980]. *Economics and Consumer Behavior*. New York: Cambridge University Press.

DeMilner, Lawrence E. [1981]. Congress of the United States, Congressional Budget Office. *Indexing with the Consumer Price Index: Problems and Alternatives*. Washington, DC: U.S. Government Printing Office (Document Number 80-400-0-81-3). [Author identified in preface]

Fisher, Franklin M. and Karl Shell [1972]. *The Economic Theory of Price Indices*. New York and London: Academic Press.

Friedman, Milton [1957]. *A Theory of the Consumption Function*. Princeton: Princeton University Press.

Gillingham, Robert F. [1974]. ''A Conceptual Framework of the Consumer Price Index''. American Statistical Association, *Proceedings of the Business and Economic Statistics Section*. Washington, DC. pp.246-52.

__________ [1983]. ''Measuring the Cost of Shelter of Homeowners: Theoretical and Empirical Considerations''. *Review of Economics and Statistics*, May 1983, Vol. LXV. pp. 254-265.

__________ and John Greenlees [1983]. ''The Incorporation of Direct Taxes in a Consumer Price Index''. *Price Level Measurement: Proceedings From a Conference Sponsored by Statistics Canada.*

Goldberger, Arthur S. and Theodore Gameletsos [1970]. ''A Cross-Country Comparison of Consumer Expenditure Patterns''. *European Economic Review* 1:357-400.

Goldfeld, Stephen M. and Van Dorn Ooms [1981]. Executive Office of the President, Council of Economic Advisers and Office of Management and Budget. 1981. *Report on Indexing Federal Programs*. Washington, DC: U.S. Government Printing Office (Document Number: 1981 0-335-818). [Authors not explicitly identified in publication]

Harberger, Arnold C., ed. [1960]. *The Demand for Durable Goods*. Chicago: The University of Chicago Press.

Kahn, Alfred E. [1980]. ''Housing Component of the Consumer Price Index''. Testimony before the Task Force on Inflation of the House Budget Committee, January 24, 1980. pp.31-64. Committee Serial No. TF-6-96-23.

Konüs, A.A. [1939]. ''The Problem of the True Index of the Cost-of-Living''. *Econometrica* 7:10-29.

Manser, Marilyn E. ''A Note on Cost-of-Living Indexes and Price Indexes for U.S. Food Consumption, 1948-1972''. BLS Working Paper 57. Washington, DC: U.S. Department of Labor, Bureau of Labor Statistics.

Muth, Richard F. [1974]. ''Moving Costs and Housing Expenditure''. *Journal of Urban Economics* 1:108-25.

Oswald, Rudolph A. [1980]. ''Housing Component of the Consumer Price Index''. Testimony before the Task Force on Inflation of the House Budget Committee, January 24, 1980. pp.64-89. Committee Serial No. TF-6-96-23.

Pencavel, John H. [1977]. ''Constant-Utility Index Numbers of Real Wages''.*American Economic Review* 67:97-100.

Phlips, Louis [1974]. *Applied Consumption Analysis*. Oxford and New York: North-Holland Publishing Company and American Elsevier Publishing Company, Inc.

Pollak, Robert A. [1971]. ''The Theory of the Cost-of-Living Index''. BLS Working Paper 11. Washington, DC: U.S. Bureau of Labour Statistics. Revised version published in the present volume.

——— [1975a]. ''The Intertemporal Cost-of Living-Index''. *Annals of Economic and Social Measurement* 4:179-95.

_________ [1975b]. ''Subindexes of the Cost-of-Living''. *International Economic Review* 16:135-150.

Price Statistics Review Committee [1961]. *The Price Statistics of the Federal Government.* Printed as: U.S. Congress, Joint Economic Committee, *Government Price Statistics, Hearings*, Part I. 87th Congress, 1st Session; also printed as: National Bureau of Economic Research, General Series Number 73.

Samuelson, Paul A. and S. Swamy [1974]. ''Invariant Economic Index Numbers and Canonical Duality: Survey and Synthesis''. *American Economic Review* 64:566-93.

Shapiro, Perry and Terence Smith [1981]. ''Preferences for nonmarket goods revealed through market demand''. *Advances in Applied Microeconomics* (JAI Press, Inc.) 1:105-122.

Statistics Canada [1982]. Price Measurement Review Program, ''Consultations Feedback Report'' (draft). Ottawa: Statistics Canada.

Triplett, Jack E. [1983a]. "An Essay on Labor Cost". In Jack E. Triplett, ed., *The Measurement of Labor Cost.* Conference on Research in Income and Wealth, Studies in Income and Wealth, Vol. 48, Chicago, Illinois: The University of Chicago Press for the National Bureau of Economic Research.

_________ [1983b]. ''Three Studies of Indexes and Indexing: A Review Essay'', in the *Journal of Money, Credit and Banking: 377-385.*

U.S. Congress, Task Force on Inflation of The House Budget Committee. *Hearings.* Jan. 24, 1980. Committee Serial No. TF-6-96-23.

COMMENTS

Gordon Schaefer
Securities Department
Bank of Canada

The paper by Jack Triplett addresses an issue often referred to at this conference. The proposition that the purpose of an index determines its manner of construction, both theoretical and practical, is indeed part of everyone's set of principles. Unfortunately, actual practice frequently does not conform to this high ideal. In his paper Triplett does us he service of discussing several types of indexes which answer different cost-of-living quesions. He then considers the circumstances under which these indexes could be used for scalation purposes in private and social contracts. This is a useful exercise and few would lisagree with the conclusion that the selection of an escalation index depends on what one s indexing and that the problems involved are often more significantly social than technical. Iowever, some of these difficulties seem to stem primarily from Triplett's discussion of ncome equity problems and are less intractable if efficiency is the objective. I will argue n this comment that in many cases, from the efficiency perspective, the problem of selecng an escalation index has already been solved by private negotiations (the market). Furermore, it will be argued that non-negotiated government escalation agreements can asonably use purchasing power or expenditure based indexes.

Private sector contracts may be viewed as freely negotiated and market determined. As ointed out by Triplett, the costs associated with negotiating contracts create an incentive r multi-year agreements. Such agreements, however, increase the risk of unanticipated anges in real rewards, thereby creating a demand for some type of insurance. This is ten provided by escalation clauses,[1] which also facilitate the next set of negotiations by stricting the size of the gap between desired and actual rewards. The problem is to select a index. For example, the appropriate index for a wage contract should result in labour reiving the value of its marginal revenue product. The correct escalation index is thus fm specific and should be based on the prices of the goods produced by the firm.[2] More gerally, escalation should be with respect to economic decision variables such as wages ad interest rates and not to total real income. (Contracts rarely guarantee factor utilizain.) From labour's perspective, the use of a firm's prices produces verifiability problems;

but also, workers will prefer an index based on their particular consumption patterns. The official Consumer Price Index is the compromise index even though the verification problem is the only one it resolves adequately. However, the inadequacies of the CPI can be corrected to some extent by the negotiation process, with its richness of formulae, caps, triggers, fold-ins, etc., which effectively produces a new escalation index. Accordingly, with this sort of market solution, the actual type of index used is not so important as long as the index is understood, verifiable, and produced with a flexibility that permits negotiators to construct sub-indexes. It is also important that the index not be perverse and that measurement techniques be in some sense correct. The recent complaints about the U.S. index referred to by Triplett were no doubt founded on a perceived perverseness of that index, i.e., its treatment of mortgage rates and housing was conceptually inappropriate.

The significance of the distinction between consumer and producer prices for escalation purposes can be illustrated using the data base of RDXF, the Bank of Canada's quarterly forecasting model.[3] For an output variable defined as gross private business product, there are corresponding factor inputs and prices. Labour's price is the average weekly wage adjusted for supplementary labour income. Chart 1 plots ''real'' wages using consumer prices (CPI) and the price of output (PGPP) from 1961 to 1981. The divergence in the mid-seventies is due primarily to strong terms of trade effects which are included in PGPP but excluded from the CPI. Clearly the ratio of the two prices has differed from unity for long periods of time and has only recently begun to narrow with the deterioration in Canada's terms of trade. These divergent movements are indicative of the type of risk that is present in a wage contract, e.g., even full indexation on the CPI in the 1970s would have resulted in workers receiving a nominal wage substantially below that implied by PGPP developments. The presence of partial indexation[4] and the use of the CPI thus created pressure for the very large catch-up settlements which occurred when contracts were renegotiated. Also, as might have been expected, the incidence of work stoppages increased substantially from 1973 to 1976.

In the public sector where **non-negotiated** escalation for taxes, pensions, etc. is the norm, the problem is that there is no obvious factor price. But it should still be possible to use, if not one price index, one concept for selecting the escalation index. First, it should be emphasized that governments need not be concerned with explicitly indexing total income,

i.e., partial income indexes are irrelevant. The social contract a government imposes on itself (or has imposed on it by the polling booth) is with respect to specific programs. If these programs constitute all of a recipient's income the government will then be fully indexing income. If not, presumably the recipient can act to protect the real value of his income from other sources. Second, contributions to the government by individuals to finance present and future income transfers are made in currency or generalized purchasing power and the transfers should be of a similar nature. The appropriate escalation index is thus a purchasing power or expenditure index. Beyond these narrow points, however, the social questions of welfare, equity, and income distribution raised by Triplett can no longer be avoided. Indeed, even the issue of the utility of escalation clauses in terms of the trade-off between real output and wage and price stability goes well beyond the question of efficiency.

Footnotes

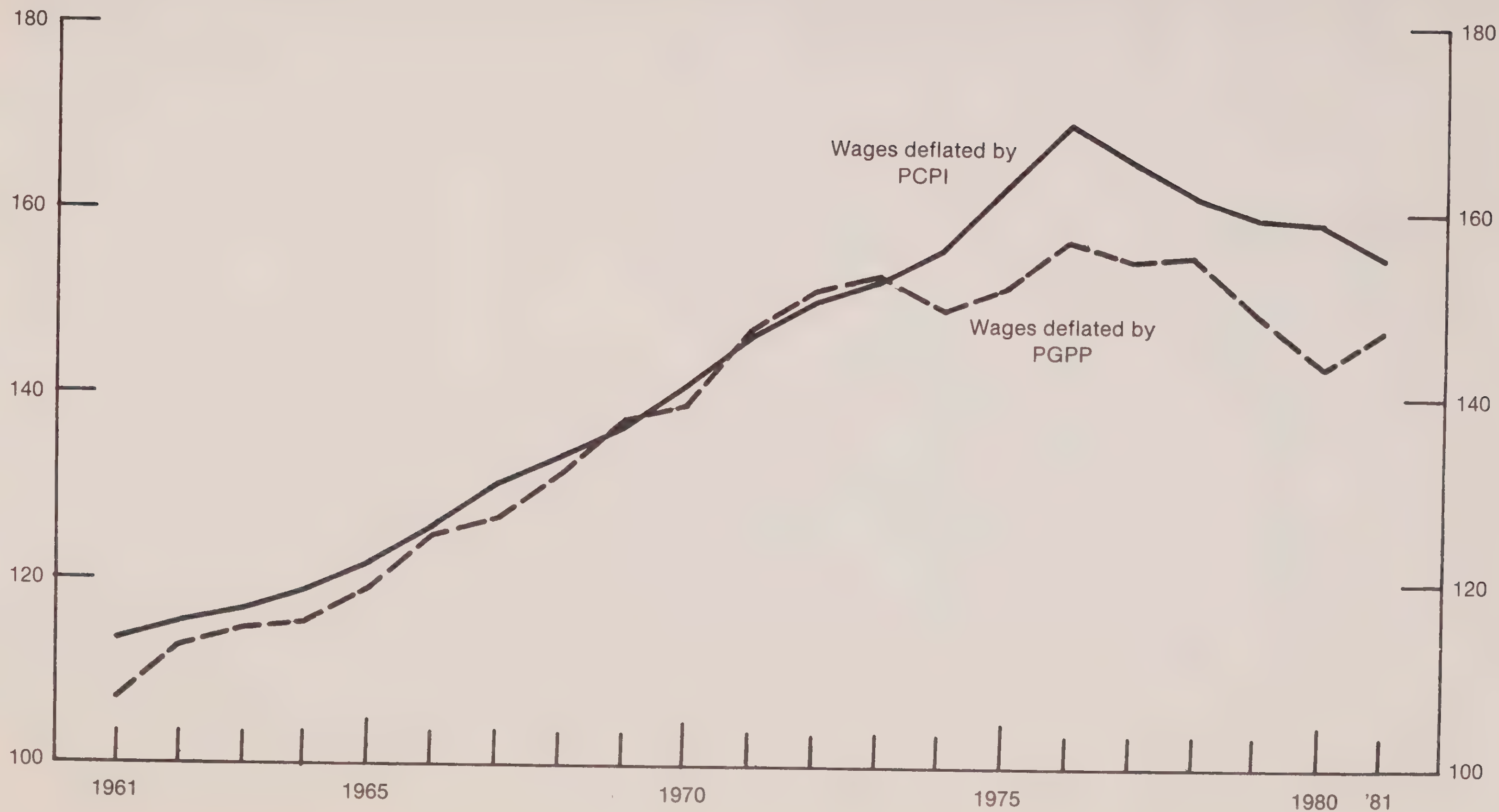
[1] See, for example, A. Blinder, ''Indexing the economy through financial intermediation'', in *Stabilization of the Domestic and International Economy*, Carnegie - Rochester Conference Series on Public Policy, pp.69-105, Vol. 5, 1978; eds., K. Brunner and A.M. Meltzer.

[2] A recent contract at Inco does specify that a wage increase will take place if the price of nickel reaches $3.20/lb.

[3] ''The Equations of RDXF'' and ''The Structure and Dynamics of RDXF'', Technical reports 25 and 26, Bank of Canada, 1982. The data used in this note are derived from the *National Expenditure and Income Accounts*.

[4] For example, the elasticity of wage increases with respect to prices in contracts with cost of living clauses seems to be between .4 and .6 and rises slightly with inflation. (See ''Technical note on cost-of-living adjustment clauses in major wage settlements'', *Bank of Canada Review*, December 1979). As inflation accelerates and as relative price movements become less significant these elasticities should approach one. The limited range of Canadian and U.S. inflation is inadequate for evaluating this hypothesis. Note also that Wilton finds an elasticity of .96 for contracts without triggers, caps, etc. (D.A. Wilton, ''An Analysis of Canadian Wage Contracts with COLA Clauses'', *Economic Council of Canada*, Discussion Paper No. 165, March 1980.)

Chart — 1
Two Measures of "Real" Wages



IMPACT OF THE CHOICE OF FORMULAE ON THE CANADIAN CONSUMER PRICE INDEX*

Pour vous fournir une version dans la langue officielle de votre choix, le texte anglais est suivi du texte français (p.512) dans cette publication.

Pierre A. Généreux
Statistics Canada

SUMMARY

One of the questions considered by Prices Division of Statistics Canada in the construction of the Consumer Price Index (CPI) is the choice of an index-number formula. There are a variety of price index formulae to choose from; they are based on different concepts, display different characteristics and potentially produce different numerical results. Differences in numerical results are yielded, in particular, by price indexes relating to different fixed baskets, a range of them being usually offered to CPI makers. This study is an attempt to quantify the differences that would result if other index-number formulae (i.e., Laspeyres, Paasche, Fisher, Chain Laspeyres, Chain Paasche and Chain Fisher) were employed for the measurement of the composite consumer price change for the period 1957 to 1978 for Canada. The study shows that, for the period 1957 to 1978, the impact of the choice of formulae on the CPI was not of considerable magnitude.

Another relevant topic dealt with by the study is the impact of the frequency of updating the CPI baskets. The study compares a Chain Laspeyres series with eight basket-updatings over the 1957-1978 period (i.e., using all available baskets from the Family Expenditure Surveys) against a Chain Laspeyres series with only one basket-updating over the 21-year period (i.e., maintaining as much as possible the same effective dates for updating the weights as the official CPI series over the period in question). The difference between these two series does not seem to demonstrate the need for more frequent updatings of the CPI baskets; such difference could be reduced by implementing the new weights in the years they refer to. Hence, what appears to be desirable is not necessarily a more frequent updating of the CPI baskets but a more timely one.

Introduction

One of the questions considered by Prices Division of Statistics Canada in the construction of the Consumer Price Index (CPI) is the choice of an index-number formula. There are a variety of price index formulae to choose from; they are based on different concepts, display different characteristics and potentially produce different numerical results. Differences in numerical results are yielded, in particular, by price indexes relating to different fixed baskets, a range of them being usually offered to CPI makers.

It seems important not only to caution CPI users that their interpretation of the composite consumer price change exhibited by the official CPI series is dependent, among other things, on the particular formula used but also to quantify the differences that would result if other index-number formulae were employed for the measurement of the composite consumer price change. This study is an attempt to quantify the above differences for the period 1957 to 1978 for Canada, with reference to the following six formulae: Laspeyres, Paasche, Fisher, Chain Laspeyres, Chain Paasche and Chain Fisher. The study shows that, for the period 1957 to 1978, the impact of the choice of formulae on the CPI was not of considerable magnitude. At the All-items level, the extreme results for 1978 were obtained using the Paasche formula (225.5) and the Laspeyres one (248.0); they differ by 22.5 index points over the 21-year period, equivalent to an average difference of 0.45% per year.

Another relevant topic dealt with by the study is the impact of the frequency of updating the CPI baskets. The study compares a Chain Laspeyres series with eight basket-updatings over the 1957-1978 period (i.e., using all available baskets from the Family Expenditure Surveys) against a Chain Laspeyres series with only one basket-updating over the 21-year period (i.e., maintaining as much as possible the same effective dates for updating the weights as the official CPI series over the period in question). At the All-items level, the 1978 results produced by the two Chain Laspeyres series on a 1957 time base are 237.9 and 244.2, respectively. A difference of 6.3 index points over the 21-year period between the two series, equivalent to an average difference of 0.12% per year, does not seem to demonstrate the need for more frequent updatings of the CPI baskets.

Theoretical Background

The quantitative proportions of commodities purchased by the population undergo shifts through time due to several factors; for instance, new products are available while some disappear from the market, tastes vary, population composition changes, real income fluctuates, substitution occurs as a result of relative price changes. Among these factors, only the latter (i.e., the price-induced substitution process) produces shifts in spending patterns that are explicitly and directly related to relative changes in prices: as a result of relative price changes, consumers will generally buy relatively less of those commodities which have become relatively more expensive, and relatively more of those commodities which have become relatively cheaper. In other words, there is usually a negative correlation between these individual commodity price changes and quantity shifts (expressed as quantity ratios) in a given time interval.

Such negative correlation makes it possible to infer the relationship between the numerical results from using, in particular, the Laspeyres (L)[1] and Paasche (P)[1] formulae. The Laspeyres formula uses the base-period basket, which is characterized by the quantitative proportions before the price-induced substitution process, whereas the Paasche uses the current-period basket, which is characterized by the quantitative proportions after the substitution. Hence, considering only the price-induced substitution, the relationship between the two formulae is as follows:

$$P^{L}_{t/0} > P^{P}_{t/0}$$

where $P^{L}_{t/0}$ is the composite Laspeyres price index, comparing prices in the period t against those in period 0,

and $P^{P}_{t/0}$ is the composite Paasche price index, comparing prices in the period t against those in period 0.

This relationship can be formally inferred from the theorem of L.V. Bortkiewicz who identified three factors that determine the direction and the magnitude of the divergence between the results produced by the Laspeyres and Paasche price indexes.[2] The three factors are:

- the dispersion of the individual commodity quantity ratios (indexes) that compare quantities in the period t against those in period 0,
- the dispersion of the individual commodity price indexes that compare prices in the period t against those in period 0, and
- the correlation between these individual commodity price indexes and quantity ratios.

The extent of the divergence is determined by the product of the three factors, but the direction is determined by the last factor as it is the only one that can be positive or negative. Considering only the price-induced substitution process, the coefficient of correlation would be negative as a rule, and consequently, the inequality $P^L_{t/0} > P^P_{t/0}$ would stand.

Assuming sustained trends in relative commodity prices, more and more shifts in the quantitative proportions will result; hence, over the long run, the cumulative effect of these shifts in the quantitative proportions will produce a significant divergence between $P^L_{t/0}$ and $P^P_{t/0}$. In the short run, without such cumulative effect, the divergence between the two price indexes will tend to be less significant.

However, the above inferred relationship is only a useful proposition which holds when considering solely the price-induced substitution process. In reality, other factors (i.e., new products, different tastes, changes in population composition, variations in real incomes) influence the shifts in the quantitative proportions over time, so that the above relationship may not occur; in fact, it is even conceivable that, for some aggregates and for some time intervals, the Laspeyres price index will be less than the Paasche.

Since the Laspeyres and Paasche formulae yield different results, several suggestions have been made to circumvent the necessity of choosing between them. With reference to fixed-basket formulae only, two approaches are widely suggested for this purpose: the Fisher[3] and the Chain formulae.[3] The Fisher formula, as a geometric mean of the Laspeyres and Paasche price indexes, is explicitly designed to yield numerical results that lie between those realized by using the Laspeyres and Paasche formulae. The Chain Laspeyres(CL), the Chain Paasche(CP) and the Chain Fisher formulae reflect the on-going shifts in the quantitative proportions and can be viewed as a product of several Laspeyres, Paasche and Fisher price indexes, respectively, computed for consecutive time sub-intervals. Assuming sustained trends in relative commodity prices, it is expected that they would lie

between the numerical results produced by using the Laspeyres and Paasche formulae;[4] further, it is expected that the Chain Laspeyres price index would be higher than the Chain Paasche one, though the divergence in their numerical results tends to be relatively small. Hence, in these circumstances, the following relationships can be established:

$$P^{L}_{t/0} > P^{CL}_{t/0} > P^{CP}_{t/0} > P^{P}_{t/0}$$

where $P^{L}_{t/0}$ and $P^{P}_{t/0}$ are defined above,

$P^{CL}_{t/0}$ is the composite Chain Laspeyres price index, comparing prices in period t against those in period 0,

and $P^{CP}_{t/0}$ is the composite Chain Paasche price index, comparing prices in period t against those in period 0.

Price Index Series Considered in this Study

The study incorporates the price index series based on the Laspeyres and Paasche formulae, among others. The choice of these two formulae is justified not only by the fact that they are the two most known index-number formulae but also because they are said to yield extreme results with reference to the other fixed-basket formulae. The study also incorporates price index series based on the Fisher and on the Chain formulae as they are two approaches widely suggested to circumvent the necessity of choosing between the Laspeyres and Paasche price indexes.

The numerical results from using the six above formulae can be compared to each other, yielding differences which are attributable to the choice of formulae. The results could also be compared against the official Canada CPI series: however, the differences in the numerical results would not be due only to the impact of the choice of formulae, but also to a number of other reasons. In the first place, the individual commodity price indexes and weight data used for the computation of the six formulae under consideration are not exactly the same as those used for the computation of the official CPI series; for instance, the weight data used for the computation of the six formulae refers to only eight cities whereas the weight data used in the official CPI construction refers to a varying number

of cities, depending on the time interval considered. In order to eliminate this source of discrepancy between the results produced by using the six formulae and the official Canada CPI series, an additional series was derived (hereafter referred to as series A) using price indexes and weights extracted from the same data bank as the one put in place for the computation of the six formulae of this study, while maintaining as much as possible the same effective dates for updating the weights as in the official Canada CPI series. Furthermore, the official Canada CPI series are, in fact, chain indexes for which the updated weights are implemented with a certain time-lag, so that the updating of the CPI weights occurs several years after the period the new weights refer to; for instance, the 1957 weights were actually implemented in January 1961, the 1967 weights in April 1973 and the 1974 ones in September 1978. In order to quantify and analyze the impact of the time-lag in implementing the updated weights on the official Canada CPI series, series A was altered by implementing the 1957, 1967 and 1974 weights in 1957, 1967 and 1974, respectively; this series is referred to as series B.

Data Sources and Operations

The individual commodity price indexes and weights used in this study refer to the lowest-level commodity groupings applied in the computation of the official Canada CPI, namely to the elementary commodity groups of the 1974 CPI classification.[5] The individual commodity weights refer to eight cities[6] only and are based on expenditure data from the Family Expenditure Surveys (FEX), available for the following years: 1957, 1959, 1962, 1964, 1967, 1969, 1972, 1974, 1976 and 1978; the expenditures were reclassified according to the 1974 CPI classification. The annual Canada commodity price indexes used were extracted from the CANSIM[7] data base which stores the commodity price indexes from the official Canada CPI program; an imputation pattern was established for the non-priced elementary commodity groups. The relevant procedures necessary to derive the price and weight data for this study as well as the operations to calculate the various price index series required a considerable amount of data research and computer processing; they are further discussed in Appendix II.

The nature of the available FEX data had some impact on the price index series computed in this study; the Chain Laspeyres, Chain Paasche and Chain Fisher indexes incorporate family expenditure data for only those years for which family expenditures are

available. Furthermore, because only annual indexes were used in the computation of this study's price index series, the exact monthly effective dates for the updatings of the baskets in series A, could not be totally respected. Series A maintains, as much as possible, the same effective dates for updating the weights; for instance, in series A, the 1967 weights were implemented in 1972 instead of April 1973 as in the official Canada CPI.[8]

Results and Their Interpretations

Following is a table which presents the 1978 indexes on a 1957 time base for the different series computed in this study and for the official Canada CPI series.

TABLE I. Comparison of 1978 Indexes
(1957 = 100)

	Official	Series A	Series B	Laspeyres[1]	Chain Laspeyres[1]	Chain Paasche[1]	Paasche[1]	Fisher[1]	Chain Fisher[1]
All-items	247.8	244.2	240.5	248.0	237.9	228.8	225.5	236.5	233.3
Food	286.1	288.5	293.6	289.3	302.6	280.6	294.4	291.8	291.4
Food for home consumption	279.5	281.9	286.7	282.5	296.6	271.5	280.0	281.2	283.8
Food away from home	N/A[2]	340.4	341.0	339.9	340.2	338.2	339.4	339.6	339.2
All-items excluding Food	234.1	230.3	225.3	232.2	221.9	216.8	212.0	221.9	219.3
Housing	250.5	251.1	244.8	252.3	237.4	233.7	229.7	240.7	235.5
Clothing	195.5	193.2	189.8	196.3	187.2	185.5	182.0	189.0	186.3
Transportation	228.1	216.2	214.0	214.1	212.8	209.5	200.3	207.1	211.1
Health and personal care	265.9	252.6	246.0	259.2	241.6	235.9	230.6	244.5	238.7
Recreation, reading and education	226.3	223.7	217.5	221.5	215.9	207.6	202.5	211.8	211.7
Tobacco and alcohol	212.7	213.1	211.0	214.2	210.0	210.5	207.7	210.9	210.2

[1]For the algebraic notation, see Appendix I.
[2]Not available on a 1957 = 100 time base since this aggregate only started as of 1961 in the official Canada CPI series.

The relationships between the above numerical results for All-items and for its major components (except for the Food component, whose relationships are discussed below) are consistent with the earlier statements; these are:

$$P^{L}_{78/57} > P^{CL}_{78/57} > P^{CP}_{78/57} > P^{P}_{78/57}$$

However, the divergence between $P^{L}_{78/57}$ and $P^{P}_{78/57}$ is less than might have been expected: the 1978 All-items indexes (1957 = 100), using the Paasche and Laspeyres formulae, equal 225.5 and 248.0, respectively, the lowest and highest results. An analysis of such divergence using the Bortkiewicz Theorem was carried out; the results show, at the All-items level:

- a large dispersion in the individual commodity quantity ratios,
- a relatively smaller dispersion in the individual commodity price indexes, and
- a weak but negative correlation between these individual commodity price indexes and quantity ratios.

The large dispersion in the individual commodity quantity ratios suggests that between 1957 and 1978 major shifts in the quantitative proportions occurred.[9] On the basis of these substantial quantity shifts, a large divergence between $P^{L}_{78/57}$ and $P^{P}_{78/57}$ could be anticipated. However, as proven by the Bortkiewicz Theorem, shifts in the quantitative proportions over time are not sufficient in themselves to provide different numerical results: to obtain significantly different results also requires a large dispersion in the individual commodity price indexes and a noticeable correlation, neither of which did occur for the period 1957 to 1978 as shown by the results of the analysis using the Bortkiewicz Theorem.

At least two reasons can be put forward in order to explain the weak correlation that occurred in spite of the effects of the price-induced substitution.

Firstly, this substitution is expected to be greatest for close substitutes (for example, between butter and margarine or between different varieties of margarine). In the computation of consumer price index series whether of the Laspeyres or Paasche type, the lowest-level aggregates considered (i.e., the elementary commodity groups) do not correspond

to narrowly-defined commodities but rather to their groupings (i.e., all varieties of margarine are grouped in the aggregate 'margarine').[10] It follows then, that the price-induced substitution may occur within those lowest-level aggregates, but it will not be reflected in the measured divergence between Laspeyres and Paasche price indexes.

Secondly, as previously stated, quantity shifts result not only from the price-induced substitution process but also from other factors such as different tastes, changes in population composition, variations in real incomes. The shifts resulting from these other factors are not necessarily negatively correlated with price changes; in fact, they may be positively correlated or not correlated at all and this may have offset some of the negative correlation which results from the price-induced substitution process. An example of positive correlation is the Food aggregate;[11] more precisely, positive correlation between price and quantity changes occurred at the level of the two components, Food for home consumption and Food away from home. This more than offsets the negative correlation that exists at the lowest-level aggregates considered within each of these two food components. The positive correlation can be traced to the greater increase in the expenditure for Food away from home, despite its faster rising prices (in comparison to the price change for Food for home consumption); such positive correlation explains why $P^{P}_{78/57}$ is greater than $P^{L}_{78/57}$ for the entire Food aggregate.

The comparison of the Chain Laspeyres (CL) series and the series A is an attempt to quantify the impact of the frequency of updating the CPI baskets. The CL series incorporates the maximum number of updatings of the baskets possible over the 1957-1978 period, this number (i.e., eight basket-updatings) being determined by the availability of family expenditure data. Series A is, in fact, a Chain Laspeyres series and maintains as much as possible the same effective dates for updating the weights as in the official CPI series (i.e., one basket-updating over the 21-year period under consideration). The 1978 CL All-items stands at 237.9 and compares to 244.2 for the 1978 All-items of series A; translated into average year-to-year indexes, the series yielded 104.2 (CL) and 104.3 (series A), an average difference of less than 0.12% per year. The need for more frequent updatings of the CPI baskets does not yet seem to be clearly demonstrated by such a small difference though there might be some other valid reasons for more frequent or more regular updatings of the CPI baskets.

The difference between the Chain Laspeyres (CL) series and series A can be further reduced by implementing the new weights in the years they refer to, as it is the case in series B. Series A and B have the same characteristics except that they implement the new weights in different periods: series A maintains as much as possible the same effective dates for the updatings of the baskets as in the official CPI series whereas series B incorporates the 1957, 1967 and 1974 baskets in 1957, 1967 and 1974, respectively. The 1978 All-items for the CL series and series B are 237.9 and 240.5, respectively, a difference smaller than the one between the CL series and series A.

The 1957-1978 period is characterized by different rates of price changes; it is legitimate to ask, then, whether the divergence between the Laspeyres and the Paasche price indexes (i.e., the extreme results over the 1957-1978 period) would be different for sub-periods which feature different rates of price changes. The entire 1957-1978 period can be decomposed into two sub-periods, 1957-1972 and 1972-1978, the latter period being characterized by faster rising prices in comparison to the former period. The various price index series considered in this study have been computed for these two sub-periods in an identical way as for the entire 1957-1978 period; the results are presented in Appendix IV. The results show that the divergences for the All-items between the Laspeyres and the Paasche series are not significantly different for the two sub-periods: the divergences between the Laspeyres and the Paasche price indexes increase from an average of 0.26% per year for the 1957-1972 period to an average of 0.35% per year for the 1972-1978 period.

Concluding Remarks

For the period 1957 to 1978, the impact of the choice of formulae on the CPI was not of considerable magnitude, as the six formulae under consideration yielded rather similar results. In particular, the divergence between $P^L_{78/57}$ and $P^P_{78/57}$ is not extremely large, despite substantial shifts in the quantitative proportions of commodities; to obtain distinct results would require also a large dispersion in the individual commodity price indexes and a noticeable correlation, neither of which did occur.

Taking into account the 1957 to 1978 experience and, in particular, the fact that the difference between series A and the Chain Laspeyres series (which incorporates all of the

possible updatings of the CPI baskets) is relatively small, the need for more frequent updatings of the CPI baskets does not yet seem to be clearly demonstrated. The difference between these two series could be considerably reduced by implementing the new weights in the years they refer to. Hence, what appears to be desirable is not necessarily a more frequent updating of the CPI baskets but a more timely one.

The conclusions of this study should not be extrapolated, without a great deal of caution, to time intervals other than the period 1957-1978, to index series other than the CPI or to other countries' experience.[12] Indeed, different economic conditions may not have the same influence on price and quantity changes and their relationships; thus one can suggest that different conditions might produce numerical results different from those displayed in the study. However, this should not be over-emphasized as shown by the computations of this study; despite faster rising prices for the 1972-1978 period in comparison to the 1957-1972 period, the divergence between the Laspeyres and the Paasche price series for the period 1972-1978 is not significantly different than the divergence between the 1957-1972 period.

Footnotes

* This study was conceived and carried out as part of the programme of the Central Research Section, Prices Division, Statistics Canada, under the supervision of B.J. Szulc, Chief of the Section. The first draft of this paper was written in 1980.

[1] For the algebraic notation, see Appendix I.

[2] See *The Consumer Price Index. Revision Based on 1974 Expenditures. Concepts and Procedures*, Statistics Canada Catalogue No. 62-546, Section 5.3, or *The Consumer Price Index Reference Paper. Concepts and Procedures. Updating Based on 1978 Expenditures*, Statistics Canada Catalogue No. 62-553, Section 8.1. For an algebraic proof of the Bortkiewicz Theorem, see *Index Numbers in Theory and Practice*, by R.D.G. Allen, Aldine Publishing Company, Chicago, 1975, pp.62-64; the theorem is also discussed in Appendix I of B.J. Szulc, ''Linking Index Number Series'', *Price Level Measurement, Proceedings from a Conference Sponsored by Statistics Canada*, 1983.

[3] For the algebraic notation, see Appendix I.

[4] See Section 8 of B.J. Szulc, ''Linking Index Number Series'', *Price Level Measurement, Proceedings from a Conference Sponsored by Statistics Canada,* 1983.

[5] See Statistics Canada Catalogue No. 62-546, *op.cit.,* Sections 2.3 and 3.2.

[6] The eight cities are: St. John's, Halifax, Montréal, Ottawa, Toronto, Winnipeg, Edmonton and Vancouver.

[7] CANSIM: Canadian Socio-Economic Information Management System. It refers to Statistics Canada's publicly accessible machine-readable data base and retrieval system.

[8] In series A, the 1957 weights are implemented in 1957 and not in January 1961 as in the official series; the 1967 weights are implemented in 1972, instead of April 1973 as in the official Canada CPI series. The 1974 series are not used in series A, since they were implemented in the official series in September 1978 only and the time period covered in this study ends in 1978.

[9] See Appendix III, where 1957 and 1978 weights are presented for All-items and its major components.

[10] See Statistics Canada Catalogue No. 62-546, *op. cit.*, Section 5.2.

[11] In analyzing the Food component, one should remember that adjustments to the total food expenditure data were carried out in this project: see Appendix II.

[12] For results pertaining to USA data, see Jack E. Triplett, ''Reconciling the CPI and the PCE Deflator'', *Monthly Labour Review*, September 1981, pp.3-15. In the article, USA differences using the Laspeyres, Chain Laspeyres and Paasche formulae are quantified for the period 1974-1980.

Appendix I

Algebraic Notation of Laspeyres, Paasche, Fisher, Chain Laspeyres, Chain Paasche and Chain Fisher Formulae[1]

Laspeyres formula: $P^{L}_{t/0} = \dfrac{\Sigma\, w_0 \cdot p_{t/0}}{\Sigma\, w_0}$

where $P^{L}_{t/0}$ is the composite Laspeyres price index comparing prices in the current period t against those in the base period 0,

w_0 is the value weight assigned to a particular elementary commodity group; this weight corresponds to the quantitative proportions of the base-period basket at base-period prices,

$p_{t/0}$ is the individual price index for a particular elementary commodity group, comparing prices in period t against those in period 0,

and Σ indicates the summation over all elementary commodity groups contained in the basket.

Paasche formula: $P^{P}_{t/0} = \dfrac{\Sigma\, w_t}{\Sigma w_t \div p_{t/0}}$

where $P^{P}_{t/0}$ is the composite Paasche price index comparing prices in the current period t against those in the base period 0,

w_t is the value weight assigned to a particular elementary commodity group; this weight corresponds to the quantitative proportions of the current period basket at current period prices,

$p_{t/0}$ and Σ are defined above.

Fisher formula: $P^F_{t/0} = \sqrt{P^L_{t/0} \cdot P^P_{t/0}}$

where $P^F_{t/0}$ is the composite Fisher price index comparing prices in the current period t against those in the base period 0.

Chain Laspeyres formula: $P^{CL}_{t/0} = P^L_{1/0} \cdot P^L_{2/1} \cdot \ldots \cdot P^L_{t/t-1}$

$$= \frac{\Sigma w_0 \cdot p_{1/0}}{\Sigma w_0} \cdot \frac{\Sigma w_1 \cdot p_{2/1}}{\Sigma w_1} \cdot \ldots \cdot \frac{\Sigma w_{t-1} \cdot p_{t/t-1}}{\Sigma w_{t-1}}$$

where $P^{CL}_{t/0}$ is the composite Chain Laspeyres price index comparing prices in the current period t against those in the base period 0,

$P^L_{1/0}$, $P^L_{2/1}$ and $P^L_{t/t-1}$ are the composite Laspeyres price indexes comparing prices in periods 1, 2 and t against those in periods 0, 1 and t-1, respectively,

w_0, w_1 and w_{t-1} are the value weights assigned to a particular elementary commodity group; these weights correspond to the quantitative proportions of the baskets of periods 0, 1 and t-1 at prices of periods 0, 1 and t-1, respectively,

$p_{1/0}$, $p_{2/1}$ and $p_{t/t-1}$ are the individual price indexes for a particular elementary commodity group, comparing prices in periods 1, 2 and t against those in periods 0, 1 and t-1, respectively,

1, 2 and t-1 indicate periods between 0 and t.

Chain Paasche formula: $P^{CP}_{t/0} = P^P_{1/0} \cdot P^P_{2/1} \cdot \ldots \cdot P^P_{t/t-1}$

$$= \frac{\Sigma w_1}{\Sigma w_1 \div p_{1/0}} \cdot \frac{\Sigma w_2}{\Sigma w_2 \div p_{2/1}} \cdot \ldots \cdot \frac{\Sigma w_t}{\Sigma w_t \div p_{t/t-1}}$$

where $P^{CP}_{t/0}$ is the composite Chain Paasche price index comparing prices in the current period t against those in the base period 0,

$P^{P}_{1/0}$, $P^{P}_{2/1}$ and $P^{P}_{t/t-1}$ are the composite Paasche price indexes comparing prices in periods 1, 2 and t against those in periods 0, 1 and t-1, respectively,

w_1, w_2 and w_t are the value weights assigned to a particular elementary commodity group; these weights correspond to the quantitative proportions of the baskets of periods 1, 2 and t at prices of periods 1, 2 and t, respectively,

$p_{1/0}$, $p_{2/1}$, $p_{t/t-1}$, 1, 2 and t-1 are defined above.

Chain Fisher formula:

$$P^{CF}_{t/0} = \sqrt{P^{L}_{1/0} \cdot P^{P}_{1/0}} \cdot \sqrt{P^{L}_{2/1} \cdot P^{P}_{2/1}} \cdot \ldots \cdot \sqrt{P^{L}_{t/t-1} \cdot P^{P}_{t/t-1}}$$

where $P^{CF}_{t/0}$ is the composite Chain Fisher price index comparing prices in the current period t against those in the base period 0.

Appendix II

Brief Explanation of the Procedures and Adjustments in Deriving the Individual Commodity Weights and Price Indexes

The individual commodity weights were extracted mainly from the 1957 to 1978 Family Expenditure Surveys (FEX) and the individual commodity price indexes, from the CANSIM data base supplemented by an imputation pattern for the non-priced commodities.

(a) Weights

The 1957 to 1978 FEX provide the necessary weight information for this project. The FEX Full-Budget Surveys, covering mainly non-food commodities, are available for 1957, 1959, 1962, 1964, 1967, 1969, 1972, 1974, 1976 and 1978; the food expenditures, derived from the FEX Food Surveys, could be obtained for 1957, 1962, 1969, 1974, 1976 and 1978 only. Before using the data from these surveys, a number of problems had to be solved; hence, it was decided that:

(i) the Canada expenditures be based on a weighted average of eight cities' expenditures; the eight cities are: St. John's, Halifax, Montréal, Ottawa, Toronto, Winnipeg, Edmonton and Vancouver. The weights used to aggregate city expenditures reflect the derived number of spending units per city in a given survey.

(ii) for 1957, 1959 and 1962, the FEX data referring to the CPI target group be used and, for the other years, the FEX data referring to the CPI extended population group be incorporated; accordingly the 1959 detailed expenditures for the CPI target group from the 1959 FEX Full-Budget Survey had to be derived.

(iii) the 1957 to 1972 aggregate expenditures for Food at home and for Food away from home reflect the aggregate expenditures derived from the relevant 1957 to 1972 Full-Budget Surveys. The 1974, 1976 and 1978 equivalents reflect the reconciled overall food expenditures obtained from the 1974, 1976 and 1978 Full-Budget and Food Surveys. The detailed food expenditures for Food at home for 1957, 1962, 1969, 1974, 1976 and 1978 reflect the distributions that characterized the

relevant Food Surveys, while the food detailed expenditures for the missing 1959, 1964, 1967 and 1972 survey years were interpolated from the detailed expenditures of the two closest Food Surveys (i.e., 1959 food expenditures were interpolated based on 1957 and 1962 data). The detailed food expenditures for Food away from home for all years in question reflect the 1974 distributions.

(iv) the weights for depreciation be estimated and the alcoholic beverages' expenditures be adjusted according to the procedures adopted for the 1974 CPI Revision.[2]

(v) the commodity expenditures from the different FEX be reclassified according to the 1974 CPI classification using, if necessary, the expenditure distributions from the closest more recent FEX (i.e., 1959 be based on 1962, whenever necessary).

(b) Price Indexes

A price index series covering the period 1957 to 1978 must be assigned to each of the approximately 700 commodities, classified according to the 1974 CPI classification by elementary commodity groups. For many commodities, the price index series are available from CANSIM in machine-readable form; for the others (i.e., non-priced commodity groups), it was decided to rely on the 1974 CPI Imputation Pattern modified to suit this study's needs.

Appendix III

Comparison of Weights Founded on the 1957 and 1978 Baskets

Major Components	Values of 1957 basket	Values of the 1978 basket	
	Expressed in 1957 prices		Expressed in 1978 prices
All-items	100.00	100.00	100.00
Food	27.71	16.38	21.38
Food for home consumption	24.41	12.41	15.41
Food away from home	3.30	3.97	5.97
All-items excluding food	72.29	83.62	78.62
Housing	33.65	34.62	35.26
Clothing	11.01	11.92	9.62
Transportation	12.93	17.61	15.64
Health and personal care	3.92	3.67	3.75
Recreation, reading and education	5.18	9.40	8.44
Tobacco and alcohol	5.60	6.40	5.91

Appendix IV

Additional Results: 1972 Indexes (1957 = 100) and 1978 Indexes (1972 = 100) for the Different Price Index Series Considered in this Study

Following are two tables which present the 1972 indexes on a 1957 time base and the 1978 indexes on a 1972 time base for the different series computed in this study and for the official Canada CPI series.

TABLE II. Comparison of 1972 Indexes
(1957 = 100)

	Official	Series A	Series B	Laspeyres	Chain Laspeyres	Chain Paasche	Paasche	Fisher	Chain Fisher
All-items	148.2	146.3	146.0	146.3	145.4	142.4	140.8	143.5	143.9
Food	148.0	146.7	148.4	146.7	150.4	145.4	146.8	146.7	147.9
Food for home consumption	143.9	142.7	143.9	142.7	145.5	139.6	140.9	141.8	142.5
Food away from home	N/A[1]	176.7	176.6	176.7	176.3	176.0	176.0	176.3	176.1
All-items excluding Food	148.1	146.1	145.1	146.1	143.6	141.5	139.2	142.6	142.5
Housing	150.6	151.2	150.6	151.2	148.5	146.3	143.9	147.5	147.4
Clothing	137.0	134.9	135.0	134.9	134.0	133.4	131.7	133.3	133.7
Transportation	144.3	137.5	136.5	137.5	135.9	134.2	132.5	135.0	135.0
Health and personal care	167.7	159.2	156.9	159.2	155.0	151.9	147.8	153.4	153.4
Recreation, reading and education	156.9	152.7	151.6	152.7	150.8	145.7	138.6	145.5	148.2
Tobacco and alcohol	140.5	142.4	141.1	142.4	140.5	140.8	139.6	141.0	140.6

[1] Not available on a 1957 = 100 time base since this aggregate only started as of 1961 in the official Canada CPI series.

TABLE III. Comparison of 1978 Indexes (1972 = 100)

	Official	Series A	Series B	Laspeyres	Chain Laspeyres	Chain Paasche	Paasche	Fisher	Chain Fisher
All-items	167.2	166.9	164.6	163.7	163.6	160.7	160.3	162.0	162.1
Food	193.3	196.7	197.9	197.4	201.2	193.0	195.9	196.6	197.1
Food for home consumption	194.3	197.5	199.2	198.5	203.8	194.5	197.3	197.9	199.1
Food away from home	187.7	192.6	193.1	192.9	193.0	192.2	192.3	192.6	192.6
All-items excluding Food	158.1	157.6	154.9	154.1	154.5	153.2	152.8	153.4	153.8
Housing	166.3	166.1	162.5	160.5	159.9	159.7	158.9	159.7	159.8
Clothing	142.7	143.2	139.8	140.4	139.7	139.1	139.2	139.8	139.4
Transportation	158.1	157.2	156.2	157.1	156.6	156.1	154.9	156.0	156.3
Health and personal care	158.6	158.7	156.8	155.9	155.9	155.3	154.8	155.3	155.6
Recreation, reading and education	144.2	146.5	143.5	141.8	143.2	142.5	143.7	142.7	142.8
Tobacco and alcohol	151.4	149.6	149.5	149.5	149.5	149.5	149.5	149.5	149.5

Footnotes

[1] For the sake of simplicity, indexes in the formulae of Appendix I are written in ratio form rather than in percentage form, even though the latter is commonly used in the publication of index numbers.

[2] See Statistics Canada Catalogue No. 62-546, *op. cit.*, Section 3.5.

INCIDENCE DU CHOIX DES FORMULES SUR L'INDICE DES PRIX À LA CONSOMMATION DU CANADA*

To provide you with a version in the official language of your choice, the French text is preceeded by the English text (p.489) in this publication.

Pierre A. Généreux
Statistique Canada

RÉSUMÉ

Parmi les questions dont doit tenir compte la Division des prix de Statistique Canada dans l'élaboration de l'Indice des prix à la consommation (IPC), il y a celle du choix de la formule qui servira au calcul de l'indice. Il existe en effet toute une gamme de formules possibles: elles sont fondées sur des concepts différents, elles ne présentent pas les mêmes caractéristiques et elles peuvent produirent des résultats numériques différents. Ceux-ci diffèrent, en particulier, lorsque les indices de prix se rapportent à des paniers fixes différents, un vaste choix de ces derniers s'offrant en général aux responsables du calcul de l'IPC. L'objet de la présente étude est de quantifier les différences qu'on observerait si on employait, pour mesurer l'indice global du changement des prix à la consommation pour les périodes allant de 1957 à 1978, d'autres formules d'indices (c'est-à-dire Laspeyres, Paasche, Fisher, Laspeyres en chaîne, Paasche en chaîne et Fisher en chaîne). L'étude démontre que l'incidence du choix des formules sur l'IPC n'est pas considérable pour la période en question.

L'étude traite aussi d'un autre aspect important, celui de la fréquence des mises à jour des paniers de l'IPC. Une comparaison est établie entre un indice en chaîne de Laspeyres comportant huit mises à jour du panier entre 1957 et 1978 (c'est-à-dire qu'on a utilisé tous les paniers établis par les Enquêtes sur les dépenses des familles) et un indice en chaîne de Laspeyres comportant seulement une mise à jour du panier en 21 ans (c'est-à-dire, en utilisant, autant que possible, les mêmes dates de mise à jour des poids que pour l'IPC officiel au cours de la période à l'étude). La différence entre ces deux séries ne semble

pas indiquer qu'il soit nécessaire d'augmenter la fréquence des mises à jour des paniers de l'IPC; on pourrait réduire cette différence en appliquant les nouveaux poids aux années mêmes auxquelles ils correspondent. Par conséquent, l'amélioration qui semble souhaitable n'est pas nécessairement l'accroissement de la fréquence des mises à jour des paniers de l'IPC, mais plutôt la réduction des délais d'introduction de ces mises à jour.

Introduction

Parmi les questions dont doit tenir compte la Division des prix de Statistique Canada dans l'élaboration de l'Indice des prix à la consommation (IPC), il y a celle du choix de la formule qui servira au calcul de l'indice. Il existe en effet toute une gamme de formules possibles; elles sont fondées sur des concepts différents, elles ne présentent pas les mêmes caractéristiques et elles peuvent produire des résultats numériques différents. Ceux-ci diffèrent, en particulier, lorsque les indices de prix se rapportent à des paniers fixes différents, un vaste choix de ces derniers s'offrant en général aux responsables du calcul de l'IPC.

Il apparaît important non seulement d'avertir les utilisateurs de l'IPC que leur interprétation de l'évolution globale des prix à la consommation présentée par l'IPC officiel dépend, entre autres choses, de la formule particulière employée, mais aussi de mesurer les différences qu'on observerait si d'autres formules d'indices étaient utilisées. L'objet de la présente étude est de quantifier ces différences pour la période allant de 1957 à 1978, au Canada, dans le cas des six formules suivantes: Laspeyres, Paasche, Fisher, Laspeyres en chaîne, Paasche en chaîne et Fisher en chaîne. L'étude démontre que l'incidence du choix des formules sur l'IPC n'est pas considérable pour la période en question. En ce qui concerne les indices d'ensemble, les résultats extrêmes pour 1978 sont ceux obtenus au moyen de la formule de Paasche (225.5) et de celle de Laspeyres (248.0); la différence est de 22.5 points sur l'ensemble de la période de 21 ans, ce qui correspond à une différence moyenne de 0.45% par année.

L'étude traite aussi d'un autre aspect important, celui de la fréquence des mises à jour des paniers de l'IPC. Une comparaison est établie entre un indice en chaîne de Laspeyres comportant huit mises à jour du panier entre 1957 et 1978 (c.-à-d. qu'on a utilisé tous les paniers disponibles d'après les Enquêtes sur les dépenses des familles) et un indice en

chaîne de Laspeyres comportant seulement une mise à jour du panier au cours de la période de 21 ans (c.-à-d., en utilisant autant que possible, les mêmes dates de mises à jour des poids que pour l'IPC officiel au cours de la période à l'étude). Au chapitre des indices d'ensemble, les résultats pour 1978 des deux indices en chaîne de Laspeyres, par rapport à l'année de base 1957, étaient de 237.9 et 244.2 respectivement. Une telle différence de 6.3 points entre les deux séries sur l'ensemble de la période, qui équivaut à une différence moyenne de 0.12% par année, ne semble pas indiquer qu'il soit nécessaire d'augmenter la fréquence des mises à jour des paniers de l'IPC.

Fondements théoriques

Les proportions quantitatives des produits achetés par la population changent avec le temps en raison de divers facteurs; par exemple, de nouveaux produits arrivent sur le marché tandis que d'autres en sont retirés, les goûts évoluent, la structure démographique se transforme, le revenu réel fluctue, ou encore il se produit une substitution par suite de variations de prix relatives. De tous ces facteurs, seul le dernier (le phénomène de substitution attribuable aux prix) entraîne des modifications des habitudes de dépenses qui sont explicitement et directement reliées aux variations relatives des prix. En effet, par suite de telles variations, les consommateurs achèteront relativement moins des produits qui sont devenus relativement plus coûteux et relativement plus de ceux qui sont devenus relativement moins coûteux. Autrement dit, il existe habituellement une corrélation négative entre les variations des prix et les variations des quantités des produits individuels (exprimées sous forme de rapports de quantités) au cours d'un intervalle de temps donné.

La présence d'une telle corrélation négative permet de prévoir la relation qui existera entre les résultats que donneront, notamment, les formules de Laspeyres (L)[1] et de Paasche (P)[1]. La formule de Laspeyres utilise le panier de la période de base, c'est-à-dire qu'elle fait intervenir les proportions quantitatives qui existaient avant que ne s'opère le processus de substitution attribuable aux prix, tandis que celle de Paasche utilise le panier de la période courante, c'est-à-dire qu'elle fait intervenir les proportions quantitatives qui existent après la substitution. Par conséquent, si l'on ne tient compte que du phénomène de substitution attribuable aux prix, la relation entre les deux formules est la suivante:

$$P^{L}_{t/0} > P^{P}_{t/0}$$

où $P^{L}_{t/0}$ est l'indice synthétique des prix de Laspeyres comparant les prix de la période t à ceux de la période 0,

et $P^{P}_{t/0}$ est l'indice synthétique des prix de Paasche comparant les prix de la période t à ceux de la période 0.

Cette relation peut être déduite théoriquement du théorème de L.V. Bortkiewicz, qui a démontré que la direction et l'ampleur de l'écart entre les résultats produits par les indices de prix de Laspeyres et de Paasche dépendent de trois facteurs[2]:

- la dispersion des rapports (indices) de quantités des produits individuels comparant les quantités de la période t à celles de la période 0,
- la dispersion des indices de prix des produits individuels comparant les prix de la période t à ceux de la période 0, et
- la corrélation entre ces indices de prix et ces rapports de quantités des produits individuels.

L'ampleur de l'écart est déterminée par le produit des trois facteurs, mais sa direction ne dépend que du dernier facteur, car lui seul peut être positif ou négatif. Si l'on ne fait intervenir que le processus de substitution attribuable aux prix, le coefficient de corrélation est normalement négatif, de sorte que l'inégalité $P^{L}_{t/0} > P^{P}_{t/0}$ s'applique.

Si les prix relatifs des produits évoluent suivant des tendances qui se maintiennent, il en résultera de plus en plus de variations des proportions quantitatives, dont l'effet cumulatif à long terme sera de produire un écart appréciable entre $P^{L}_{t/0}$ et $P^{P}_{t/0}$. À court terme, sans l'action de cet effet cumulatif, les deux indices de prix présenteront généralement un écart moins significatif.

Toutefois, l'inégalité établie ci-dessus n'est qu'une relation utile qui s'applique lorsqu'on considère uniquement le phénomène de substitution attribuable aux prix. En réalité, d'autres facteurs (par ex. l'apparition de nouveaux produits, l'évolution des goûts, les changements démographiques, les fluctuations des revenus réels) influent sur les variations des proportions quantitatives qui surviennent avec le temps, de sorte que la relation ci-dessus peut très bien ne plus s'appliquer; en fait, on peut même concevoir que pour certains agrégats et certains intervalles de temps, l'indice de prix de Laspeyres sera inférieur à celui de Paasche.

Comme les formules de Laspeyres et de Paasche donnent des résultats différents, on a proposé plusieurs solutions dans le but d'éviter d'avoir à choisir entre les deux. Dans le contexte des indices à panier fixe, deux formules sont fréquemment suggérées: l'indice de Fisher[3] et l'indice en chaîne[3]. La formule de Fisher, qui consiste à faire la moyenne géométrique des indices de prix de Laspeyres et de Paasche, vise explicitement à donner des résultats numériques intermédiaires entre ceux produits par les formules de Laspeyres et de Paasche. Quant aux indices en chaîne de Laspeyres (CL), de Paasche (CP) et de Fisher, ils tiennent compte de variations continues des proportions quantitatives et peuvent être considérés comme le produit de plusieurs indices de Laspeyres, de Paasche et de Fisher respectivement, calculés pour des sous-intervalles de temps consécutifs. Si les prix relatifs des produits évoluent suivant des tendances soutenues, ces indices en chaîne devraient normalement donner des résultats numériques se situant entre ceux des formules de Laspeyres et de Paasche[4]; en outre, l'indice en chaîne de Laspeyres devrait normalement être supérieur à l'indice en chaîne de Paasche, l'écart étant cependant relativement faible. Il est donc possible, dans ces conditions, d'établir les relations suivantes:

$$P^{L}_{t/0} > P^{CL}_{t/0} > P^{CP}_{t/0} > P^{P}_{t/0}$$

où $P^{L}_{t/0}$ et $P^{P}_{t/0}$ sont définis ci-dessus,

$P^{CL}_{t/0}$ est l'indice synthétique en chaîne de Laspeyres comparant les prix de la période t à ceux de la période 0,

et $P^{CP}_{t/0}$ est l'indice synthétique en chaîne de Paasche comparant les prix de la période t à ceux de la période 0.

Indices de prix visés par l'étude

L'étude porte sur des indices de prix fondés notamment sur les formules de Laspeyres et de Paasche. Il était normal de choisir ces deux formules non seulement parce qu'elles sont les deux plus connues, mais aussi parce qu'elles sont censées donner des résultats extrêmes comparativement aux autres formules à panier fixe. L'étude s'intéresse également

à l'indice de Fisher et aux indices en chaîne, car il s'agit là de deux solutions qu'on suggère fréquemment pour éviter d'avoir à choisir entre les indices de prix de Laspeyres et de Paasche.

En comparant les résultats obtenus au moyen de chacune des six formules mentionnées, on peut constater les différences qui résultent du choix exercé. On pourrait aussi comparer les résultats obtenus aux séries de l'IPC officiel du Canada, mais les différences observées ne pourraient pas dans un tel cas être attribuées uniquement au choix d'une formule plutôt qu'une autre, car un certain nombre d'autres facteurs interviendraient. En premier lieu, les indices de prix des produits individuels et les poids utilisés dans le calcul des six formules à l'étude ne sont pas exactement les mêmes que ceux qui servent au calcul de l'IPC officiel; ainsi, les poids utilisés dans les six formules sont fondés sur les données de huit villes seulement, tandis que les poids de l'IPC officiel sont établis en fonction d'un nombre variable de villes, tout dépendant de l'intervalle de temps considéré. Afin d'éliminer cette cause d'écart entre les résultats obtenus au moyen des six formules et ceux de l'IPC officiel, on a calculé une autre série (désignée ci-après la série A) à l'aide des indices de prix et des poids tirés de la même banque de données que celle qui a servi au calcul des six formules de l'étude, mais en utilisant dans la mesure du possible les mêmes dates de mise à jour des poids que l'IPC officiel. En second lieu, il faut remarquer que l'IPC officiel est un indice en chaîne dont la mise à jour des poids se fait avec un certain décalage. La mise à jour survient en effet plusieurs années après la période à laquelle les nouveaux poids se rapportent; par exemple, les poids de 1957 ont été réellement introduits en janvier 1961, ceux de 1967, en avril 1973 et ceux de 1974, en septembre 1978. Pour qu'il soit possible de mesurer et d'analyser l'effet de cette introduction décalée des poids révisés sur l'IPC officiel, on a modifié la série A en appliquant les poids de 1957, 1967 et 1974 à ces mêmes années; cette dernière série est appelée la série B.

Sources de données et opérations

Les indices de prix et les poids des produits individuels utilisés dans cette étude sont ceux des groupes de produits du plus bas niveau qui servent au calcul de l'IPC officiel du Canada, c'est-à-dire les groupes élémentaires de produits de la classification de l'IPC de 1974[5]. Les poids des produits individuels sont établis en fonction des données de huit villes[6] seulement et sont fondés sur les chiffres des Enquêtes sur les dépenses des familles disponibles

pour les années suivantes: 1957, 1959, 1962, 1964, 1967, 1969, 1972, 1974, 1976 et 1978; les dépenses ont été reclassées en fonction de la classification de l'IPC de 1974. Les indices de prix annuels des produits pour le Canada ont été tirés de la base de données CANSIM[7], dans laquelle on retrouve les indices de prix des produits qui servent au calcul de l'IPC officiel; une règle d'imputation a été élaborée pour tenir compte des groupes élémentaires de produits non observés. Les méthodes dont on s'est servi dans la présente étude pour obtenir les prix et les poids, ainsi que les opérations qui ont permis de calculer les divers indices de prix, ont exigé un travail considérable de recherche et de traitement informatique; on trouvera des détails supplémentaires à ce sujet à l'annexe II.

La nature des données pouvant être tirées des Enquêtes sur les dépenses des familles a eu un effet sur les indices de prix calculés dans l'étude; les indices en chaîne de Laspeyres, de Paasche et de Fisher ne tiennent compte des données sur les dépenses des familles que pour les années où ces données sont disponibles. En outre, parce qu'on ne s'est servi que d'indices annuels dans le calcul des séries d'indices de prix de la présente étude, il n'a pas été possible de respecter entièrement dans la mise à jour des paniers de la série A les dates exactes où étaient survenues les mises à jour réelles. On a toutefois retenu pour la mise à jour des poids de la série A les dates qui se rapprochaient le plus des dates réelles; par exemple, on a considéré que les poids de 1967 étaient appliqués en 1972, plutôt qu'en avril 1973 comme cela a été fait pour l'IPC officiel[8].

Résultats et interprétation

Le tableau ci-dessous présente les indices de 1978, par rapport à l'année de base 1957, des différentes séries calculées dans le cadre de cette étude, ainsi que celles de l'IPC officiel du Canada.

TABLEAU 1. Comparaison des indices de 1978
(1957 = 100)

	Série officielle	Série A	Série B	Laspeyres[1]	Laspeyres en chaîne[1]	Paasche en chaîne[1]	Paasche[1]	Fisher[1]	Fisher en chaîne[1]
Indice d'ensemble	247.8	244.2	240.5	248.0	237.9	228.8	225.5	236.5	233.3
Aliments	286.1	288.5	293.6	289.3	302.6	280.6	294.4	291.8	291.4
Aliments consommés à la maison	279.5	281.9	286.7	282.5	296.6	271.5	280.0	281.2	283.8
Repas pris à l'extérieur	N/D[2]	340.4	341.0	339.9	340.2	338.2	339.4	339.6	339.2
Ensemble sans les aliments	234.1	230.3	225.3	232.2	221.9	216.8	212.0	221.9	219.3
Habitation	250.5	251.1	244.8	252.3	237.4	233.7	229.7	240.7	235.5
Habillement	195.5	193.2	189.8	196.3	187.2	185.5	182.0	189.0	186.3
Transports	228.1	216.2	214.0	214.1	212.8	209.5	200.3	207.1	211.1
Santé et soins personnels	265.9	252.6	246.0	259.2	241.6	235.9	230.6	244.5	238.7
Loisirs, lecture et formation	226.3	223.7	217.5	221.5	215.9	207.6	202.5	211.8	211.7
Tabacs et boissons alcoolisées	212.7	213.1	211.0	214.2	210.0	210.5	207.7	210.9	210.2

[1] Voir l'annexe I pour les expressions algébriques.
[2] Chiffre non disponible sur la base 1957 = 100, cet agrégat n'ayant été inclus qu'en 1961 dans l'IPC officiel du Canada.

L'examen des données du tableau révèle que le lien entre les indices, pour ce qui est de l'indice d'ensemble et de ses composantes principales (à l'exception de la composante ''aliments'', dont les liens seront étudiés ci-dessous), est conforme à la relation que nous avons établie précédemment, c'est-à-dire:

$$P^{L}_{78/57} > P^{CL}_{78/57} > P^{CP}_{78/57} > P^{P}_{78/57}$$

Toutefois, l'écart entre $P^{L}_{78/57}$ et $P^{P}_{78/57}$ n'est pas aussi grand qu'on aurait pu le prévoir: dans le cas de l'indice d'ensemble de 1978 (1957 = 100), les formules de Paasche et de Laspeyres donnent respectivement des valeurs de 225.5 et 248.0, qui représentent en fait les deux extrêmes. Une analyse de cet écart à l'aide du théorème de Bortkiewicz a été effectuée, et on a constaté, au chapitre de l'indice d'ensemble:

- une grande dispersion des rapports de quantités des produits individuels,
- une dispersion relativement plus faible des indices de prix des produits individuels, et
- une corrélation faible, mais négative, entre les indices de prix et les rapports de quantités des produits individuels.

La grande dispersion des rapports de quantités des produits individuels semble indiquer que des variations importantes des proportions quantitatives sont survenues entre 1957 et 1978[9]. Compte tenu de ces variations, on pourrait s'attendre à un écart important entre $P^{L}_{78/57}$ et $P^{P}_{78/57}$. Toutefois, en vertu du théorème de Bortkiewicz, les variations dans le temps des proportions quantitatives ne suffisent pas à elles seules à produire des écarts; pour que la différence soit significative, il faut aussi qu'il y ait une grande dispersion des indices de prix des produits individuels ainsi qu'une corrélation notable, mais on n'observe ni l'une ni l'autre de ces conditions sur la période allant de 1957 à 1978, comme le démontrent les résultats de l'analyse effectuée en vertu du théorème de Bortkiewicz.

On peut invoquer au moins deux raisons pour expliquer la faible corrélation observée en dépit du phénomène de substitution attribuable aux prix.

En premier lieu, le processus de substitution est probablement plus prononcé dans le cas de substituts voisins (par exemple, entre le beurre et la margarine ou entre différentes

sortes de margarine). Or, les agrégats du niveau le plus bas qui entrent dans le calcul des indices des prix à la consommation, qu'ils soient de type Laspeyres ou de type Paasche, sont les groupes élémentaires de produits, qui ne correspondent pas à des produits bien définis mais plutôt à des ensembles de ces derniers (par ex. toutes les sortes de margarine sont regroupées dans l'agrégat ''margarine'')[10]. Par conséquent, s'il se produit une substitution attribuable aux prix à l'intérieur de ces agrégats du niveau le plus bas, son effet n'apparaîtra pas dans l'écart mesuré entre les indices de prix de Laspeyres et de Paasche.

En deuxième lieu, comme nous l'avons dit plus haut, les variations des quantités ne sont pas le résultat du seul processus de substitution attribuable aux prix; ils dépendent aussi d'autres facteurs, par exemple l'évolution des goûts, la transformation de la structure démographique ou les fluctuations des revenus réels. Les variations attribuables à ces autres facteurs ne présentent pas nécessairement une corrélation négative avec les variations de prix; en fait, elles peuvent présenter une corrélation positive ou pas de corrélation du tout, ce qui peut avoir pour effet d'annuler une partie de la corrélation négative qu'entraîne le processus de substitution attribuable aux prix. On trouve un exemple de corrélation positive en examinant l'agrégat des aliments[11]; en fait, on observe une corrélation positive entre les variations de prix et de quantités au niveau des deux composantes de cet agrégat, soit les aliments consommés à la maison et les repas pris à l'extérieur. Une telle corrélation fait plus qu'annuler la corrélation négative qui existe à l'échelle des agrégats du niveau le plus bas qu'englobent chacune de ces deux composantes. La corrélation positive s'explique par la plus grande augmentation des dépenses au titre des repas pris à l'extérieur malgré la progression plus rapide des prix de ces derniers (comparativement à la variation des prix des aliments consommés à la maison); c'est la raison pour laquelle $P^{P}_{78/57}$ est supérieur à $P^{L}_{78/57}$ pour l'agrégat total des aliments.

En comparant la série en chaîne de Laspeyres (CL) et la série A, on voulait mesurer l'effet de la fréquence des mises à jour des paniers de l'IPC. La série CL fait intervenir le plus grand nombre possible de mises à jour des paniers sur la période allant de 1957 à 1978 (soit huit mises à jour), compte tenu des données disponibles sur les dépenses des familles. La série A, pour sa part, est en fait une série en chaîne de Laspeyres, mais elle comporte le même nombre de mises à jour que la série de l'IPC officiel (c.-à-d. une mise à jour du panier au cours de la période de 21 ans considérée). L'indice d'ensemble CL de 1978 se situe à 237.9, comparativement à 244.2 pour l'indice d'ensemble de la série A

de 1978. Sous forme d'indices moyens d'une année sur l'autre, les valeurs sont de 104.2 (CL) et 104.3 (série A) et présentent donc une différence moyenne de moins de 0.12 % par année. Une différence si faible ne semble donc pas justifier à elle seule le besoin de mises à jour plus fréquentes des paniers de l'IPC, mais il existe peut-être d'autres bonnes raisons d'effectuer ces mises à jour de façon plus fréquente ou plus régulière.

Il est possible de réduire encore plus la différence entre la série en chaîne de Laspeyres (CL) et la série A en appliquant les nouveaux poids à l'année même à laquelle ils se rapportent, ce qu'on a fait pour la série B. La série A et la série B présentent des caractéristiques identiques, sauf que les nouveaux poids ne sont pas introduits aux mêmes périodes pour l'une et l'autre: les dates des mises à jour de la série A correspondent le plus possible aux dates des mises à jour de la série de l'IPC officiel, tandis que les mises à jour de 1957, 1967 et 1974 sont appliquées ces mêmes années dans le cas de la série B. Les indices d'ensemble de 1978 de la série CL et de la série B sont respectivement de 237.9 et 240.5 et présentent donc une différence inférieure à celle qu'on a observée entre la série CL et la série A.

La période de 1957 à 1978 se caractérise par des taux différents de variations des prix; il est donc légitime de se demander, alors, si l'écart entre les indices de prix établis selon les formules de Laspeyres et de Paasche (c'est-à-dire les résultats extrêmes pour la période 1957-1978) serait différent pour des sous-périodes affichant des taux de variation des prix différents. La période 1957-1978 peut être divisée en deux sous-périodes, 1957-1972 et 1972-1978, la seconde se caractérisant par une hausse des prix plus rapide que la première. Les diverses séries d'indices de prix analysées dans la présente étude ont été calculées pour les deux sous-périodes de la même façon que pour la période 1957-1978: les résultats sont présentés à l'annexe IV. Les résultats montrent que l'écart entre les indices d'ensemble de Laspeyres et de Paasche n'est pas significativement différent pour les deux sous-périodes: cet écart augmente d'une moyenne de 0.26% par année pour la période 1957-1972 à une moyenne de 0.35% par an pour la période 1972-1978.

Conclusion

Pour la période allant de 1957 à 1978, l'incidence du choix des formules sur l'IPC n'est pas considérable, comme le démontrent les résultats assez semblables qu'ont donnés les six formules à l'étude. En particulier, l'écart entre $P^{L}_{78/57}$ et $P^{P}_{78/57}$ n'est pas très élevé,

bien que des variations considérables des proportions quantitatives des produits soient survenues au cours de la période; il aurait fallu, pour que l'écart soit plus prononcé, que les indices de prix des produits individuels présentent aussi une grande dispersion et qu'une corrélation notable soit observée, mais ni l'une ni l'autre de ces conditions n'est remplie.

Compte tenu des résultats obtenus pour la période à l'étude, et en particulier, de la différence relativement faible entre la série A et la série en chaîne de Laspeyres (dans laquelle interviennent toutes les mises à jour possibles des paniers de l'IPC), il n'apparaît pas évident que des mises à jour plus fréquentes des paniers de l'IPC soient nécessaires. Par ailleurs, on peut réduire considérablement la différence entre les deux séries en appliquant les nouveaux poids aux années mêmes auxquelles ils correspondent. Par conséquent, l'amélioration qui semble souhaitable n'est pas nécessairement l'accroissement de la fréquence des mises à jour des paniers de l'IPC, mais plutôt la réduction des délais d'introduction de ces mises à jour.

On ne devrait pas, sans prendre d'infinies précautions, étendre par extrapolation les résultats de la présente étude à d'autres intervalles de temps que la période de 1957 à 1978, à d'autres indices que l'IPC ou à l'expérience d'autres pays[12]. En effet, des conditions économiques différentes n'auront pas nécessairement le même effet sur les variations des prix et des quantités, ainsi que sur les relations entre ces dernières; on peut donc supposer que l'incidence sur les résultats pourra différer de celle que nous avons observée ici. Cependant, comme l'indiquent nos résultats, il ne faut pas exagérer l'importance de ce facteur: malgré la hausse plus rapide des prix pour la période 1972-1978 par rapport à la période 1957-1972, l'écart entre les indices de Laspeyres et de Paasche pour la période 1972-1978 n'est pas significativement différent de leur écart pour la période 1957-1972.

Renvois

* Cette étude a été conçue et réalisée dans le cadre du programme de la Section centrale de recherche, Division des prix, Statistique Canada, sous la supervision de B.J. Szulc, chef de la section. La première ébauche de ce document a été rédigée en 1980.

[1] Voir l'annexe I pour les expressions algébriques.

[2] Voir *L'indice des prix à la consommation. Révision fondée sur les dépenses de 1974. Concepts et procédés*, Statistique Canada, n° 62-546 au catalogue, section 5.3, ou *Document de référence de l'indice des prix à la consommation. Concepts et procédés. Mise à jour fondée sur les dépenses de 1978*, Statistique Canada, n° 62-553 au catalogue, section 8.1. Pour une preuve algébrique du théorème de Bortkiewicz, voir *Index Numbers in Theory and Practice*, par R.D.G. Allen, Aldine Publishing Company, Chicago, 1975, pages 62-64; on discute également le théroème dans B.J. Szulc, "Enchaînement des indices de prix", *La Mesure du Niveau des Prix: Actes du Colloque tenus sous l'égide de Statistique Canada,* 1983 (voir l'annexe I).

[3] Voir l'annexe I pour les expressions algébriques.

[4] Voir B.J. Szulc, "Enchaînement des indices de prix" *La Mesure du Niveau des Prix: Actes du Colloque tenu sous l'égide de Statistique Canada,* 1983 (voir section 8).

[5] Voir Statistique Canada, n° 62-546 au catalogue, *op.cit.*, sections 2.3 et 3.2.

[6] St. John's (T.-N.), Halifax, Montréal, Ottawa, Toronto, Winnipeg, Edmonton et Vancouver.

[7] CANSIM: Système canadien d'information socio-économique. Il s'agit du système d'extraction et de la base de données sous forme lisible par machine de Statistique Canada, qui sont accessibles au public.

[8] Dans la série A, les poids de 1957 sont appliqués en 1957 et non pas en janvier 1961 comme dans la série officielle; les poids de 1967 sont appliqués en 1972, et non pas en avril 1973 comme dans la série officielle. Quant aux poids de 1974, ils ne sont pas utilisés dans la série A, puisqu'ils n'ont été introduits dans la série officielle qu'en septembre 1978, et que la période visée par l'étude se termine en 1978.

[9] Voir l'annexe III, où les poids de 1957 et de 1978 sont présentés pour l'indice d'ensemble et ses principales composantes.

[10] Voir Statistique Canada, n° 62-546 au catalogue, *op. cit.,* section 5.2.

[11] Dans l'analyse de la composante des aliments, il ne faut pas perdre de vue que des ajustements ont été apportés aux données sur les dépenses totales consacrées aux aliments dans le cadre du présent projet; voir l'annexe II.

[12] Pour des résultats concernant des données sur les É.-U., voir Jack E. Triplett, *"Reconciling the CPI and the PCE Deflator", Monthly Labour Review*, septembre 1981, pp.3-15. Dans cet article, les écarts des données sur les É.-U. obtenus avec les formules de Laspeyres, de Laspeyres en chaîne et de Paasche sont quantifiés pour la période 1974-1980.

Annexe I

Expressions algébriques des formules de Laspeyres, de Paasche, de Fisher, de Laspeyres en chaîne, de Paasche en chaîne et de Fisher en chaîne[1]

Formule de Laspeyres: $P^{L}_{t/0} = \dfrac{\Sigma w_0 \cdot p_{t/0}}{\Sigma w_0}$

où $P^{L}_{t/0}$ est l'indice des prix global de Laspeyres où l'on compare les prix de la période courante t à ceux de la période de base 0,

w_0 est le poids attribué à un groupe élémentaire de produits particulier; ce poids correspond aux proportions quantitatives du panier de la période de base, aux prix de la période de base,

$p_{t/0}$ est l'indice de prix individuel pour un groupe élémentaire de produits particulier, où l'on compare les prix de la période t à ceux de la période 0,

et Σ indique la somme de tous les groupes élémentaires de produits contenus dans le panier.

Formule de Paasche: $P^{P}_{t/0} = \dfrac{\Sigma\, w_t}{\Sigma w_t \div p_{t/0}}$

où $P^{P}_{t/0}$ est l'indice des prix global de Paasche où l'on compare les prix de la période courante t à ceux de la période de base 0,

w_t est le poids attribué à un groupe élémentaire de produits particulier; ce poids correspond aux proportions quantitatives du panier de la période courante, aux prix de la période courante,

$p_{t/0}$ et Σ sont définis plus haut.

Formule de Fisher: $P^{F}_{t/0} = \sqrt{P^{L}_{t/0} \cdot P^{P}_{t/0}}$

où $P^F_{t/0}$ est l'indice des prix global de Fisher où l'on compare les prix de la période courante t à ceux de la période de base 0.

Formule de Laspeyres en chaîne: $P^{CL}_{t/0} = P^L_{1/0} \cdot P^L_{2/1} \cdot \dots \cdot P^L_{t/t\text{-}1}$

$$= \frac{\Sigma w_0 \cdot p_{1/0}}{\Sigma w_0} \cdot \frac{\Sigma w_1 \cdot p_{2/1}}{\Sigma w_1} \cdot \dots \cdot \frac{\Sigma w_{t\text{-}1} \cdot p_{t/t\text{-}1}}{\Sigma w_{t\text{-}1}}$$

où $P^{CL}_{t/0}$ est l'indice des prix global de Laspeyres en chaîne où l'on compare les prix de la période courante t à ceux de la période de base 0,

$P^L_{1/0}$, $P^L_{2/1}$ et $P^L_{t/t\text{-}1}$ sont les indices de prix globaux de Laspeyres où l'on compare les prix des périodes 1, 2 et t à ceux des périodes 0, 1 et t-1, respectivement,

w_0, w_1 et $w_{t\text{-}1}$ sont les poids attribués à un groupe élémentaire de produits particulier; ces poids correspondent aux proportions quantitatives des paniers des périodes 0, 1 et t-1, aux prix des périodes 0, 1 et t-1, respectivement,

$p_{1/0}$, $p_{2/1}$ et $p_{t/t\text{-}1}$ sont les indices de prix individuels d'un groupe élémentaire de produits particulier où l'on compare les prix des périodes 1, 2 et t à ceux des périodes 0, 1 et t-1, respectivement,

1, 2 et t-1 indiquent des périodes entre 0 et t.

Formule de Paasche en chaîne: $P^{CP}_{t/0} = P^P_{1/0} \cdot P^P_{2/1} \cdot \dots \cdot P^P_{t/t\text{-}1}$

$$= \frac{\Sigma w_1}{\Sigma w_1 \div p_{1/0}} \cdot \frac{\Sigma w_2}{\Sigma w_2 \div p_{2/1}} \cdot \dots \cdot \frac{\Sigma w_t}{\Sigma w_t \div p_{t/t\text{-}1}}$$

où $P^{CP}_{t/0}$ est l'indice des prix global de Paasche en chaîne où l'on compare les prix de la période courante t à ceux de la période de base 0,

$P^P_{1/0}$, $P^P_{2/1}$ et $P^P_{t/t\text{-}1}$ sont des indices de prix globaux de Paasche où l'on compare les prix des périodes 1, 2 et t à ceux des périodes 0, 1 et t-1, respectivement,

w_1, w_2 et w_t sont les poids attribués à un groupe élémentaire de produits particulier; ces poids correspondent aux proportions quantitatives des paniers des périodes 1, 2 et t, aux prix des périodes 1, 2 et t, respectivement,

$p_{1/0}$, $p_{2/1}$, $p_{t/t\text{-}1}$, 1, 2 et t-1 sont définis plus haut.

Formule de Fisher en chaîne:

$$P^{CF}_{t/0} = \sqrt{P^{L}_{1/0} \cdot P^{P}_{1/0}} \cdot \sqrt{P^{L}_{2/1} \cdot P^{P}_{2/1}} \cdot \dots \cdot \sqrt{P^{L}_{t/t\text{-}1} \cdot P^{P}_{t/t\text{-}1}}$$

où $P^{CF}_{t/0}$ est l'indice des prix global de Fisher en chaîne où l'on compare les prix de la période courante t à ceux de la période de base 0.

Annexe II

Aperçu des méthodes de calcul et d'ajustement des poids et des indices de prix des produits individuels

Les poids des produits individuels ont été tirés des données des Enquêtes sur les dépenses des familles pour la période allant de 1957 à 1978. Les indices de prix des produits individuels, quant à eux, ont été tirés de la base de données CANSIM; une règle d'imputation a été élaborée pour les produits non observés.

(a) Poids

Les données des Enquêtes sur les dépenses des familles pour la période de 1957 à 1978 ont constitué la source à laquelle on a puisé l'information nécessaire à l'établissement des poids. On dispose d'Enquêtes sur les dépenses totales des familles, qui portent surtout sur des produits non alimentaires, pour les années 1957, 1959, 1962, 1964, 1967, 1969, 1972, 1974, 1976 et 1978; en ce qui concerne les Enquêtes sur les dépenses alimentaires, il existe des données pour 1957, 1962, 1969, 1974, 1976 et 1978 seulement. Avant d'utiliser les données de ces Enquêtes, il restait un certain nombre de problèmes à résoudre et les règles suivantes ont été adoptées:

(i) les dépenses pour l'ensemble du Canada ont été établies en fonction d'une moyenne pondérée des dépenses de huit villes (St-John's (T.-N.), Halifax, Montréal, Ottawa, Toronto, Winnipeg, Edmonton et Vancouver). Les poids utilisés pour faire la somme des dépenses des différentes villes tenaient compte du nombre d'unités de dépense par ville pour chaque enquête;

(ii) pour les années 1957, 1959 et 1962, les données sur les dépenses des familles se rapportant au groupe-cible de l'IPC ont été utilisées, tandis que pour les autres années, on s'est servi des données sur les dépenses des familles se rapportant à la population élargie de l'IPC; pour 1959, il a donc fallu calculer à partir de l'Enquête sur les dépenses totales des familles, les dépenses détaillées en ce qui concerne le groupe-cible de l'IPC;

(iii) pour les années 1957 à 1972, les dépenses globales au titre des aliments consommés à la maison et des repas pris à l'extérieur ont été établies à partir des dépenses globales observées par les Enquêtes sur les dépenses totales des familles des années correspondantes; en ce qui concerne les dépenses équivalentes de 1974, 1976 et 1978, les chiffres sont le résultat d'une conciliation entre les dépenses alimentaires globales fournies respectivement par les Enquêtes sur les dépenses totales et les Enquêtes sur les dépenses alimentaires de 1974, 1976 et 1978; les dépenses alimentaires détaillées au titre des aliments consommés à la maison pour les années 1957, 1962, 1969, 1974, 1976 et 1978 ont été établies d'après les Enquêtes sur les dépenses alimentaires des années correspondantes, tandis que les dépenses alimentaires détaillées des autres années d'enquête, soit 1959, 1964, 1967 et 1972, ont été obtenues par interpolation à partir des dépenses détaillées des deux enquêtes sur les dépenses alimentaires les plus proches (par ex., les dépenses alimentaires de 1959 ont été établies par interpolation des dépenses de 1957 et de 1962). Les dépenses alimentaires détaillées au titre des repas pris à l'extérieur pour toutes les années considérées ont été tirées des distributions de 1974.

(iv) les poids relatifs à la dépréciation ont été estimés selon la méthode adoptée pour la révision de l'IPC de 1974; on a ajusté les dépenses au titre des boissons alcoolisées à partir de la méthode adoptée pour la révision de 1974.[2]

(v) les dépenses par produit selon les différentes Enquêtes sur les dépenses des familles ont été reclassées suivant la classification de l'IPC de 1974; on a utilisé au besoin les distributions des dépenses de l'Enquête sur les dépenses des familles qui suivait immédiatement l'année considérée (par ex., les dépenses de 1959 ont été basées, chaque fois que la chose était nécessaire, sur les données de 1962).

(b) Indices de prix

Un indice de prix portant sur la période allant de 1957 à 1978 doit être attribué à chacun des produits, au nombre d'environ 700, répartis selon la classification de l'IPC de 1974 en groupes élémentaires de produits. Pour de nombreux produits, les indices de prix peuvent être obtenus du système CANSIM sous forme lisible par machine; en ce qui

concerne les autres (c.-à-d. les groupes de produits non observés), on a décidé d'appliquer la règle d'imputation de l'IPC de 1974, en y apportant certaines modifications en fonction des besoins de l'étude.

Annexe III

Comparaison des poids fondés sur les paniers de 1957 et 1978

Composantes principales	Valeurs du panier de 1957	Valeurs du panier de 1978	
	Exprimées selon les prix de 1957		Exprimées selon les prix de 1978
Ensemble	100.00	100.00	100.00
Aliments	27.71	16.38	21.38
Aliments consommés à la maison	24.41	12.41	15.41
Repais pris à l'extérieur	3.30	3.97	5.97
Ensemble sans les aliments	72.29	83.62	78.62
Habitation	33.65	34.62	35.26
Habillement	11.01	11.92	9.62
Transports	12.93	17.61	15.64
Santé et soins personnels	3.92	3.67	3.75
Loisirs, lecture et formation	5.18	9.40	8.44
Tabacs et boissons alcoolisées	5.60	6.40	5.91

Annexe IV

Résultats supplémentaires: indices de 1972 (1957 = 100) et de 1978 (1972 = 100) pour les diverses séries d'indices de prix visées par cette étude

On trouvera aux pages suivantes deux tableaux présentant les indices de 1972 sur la base de 1957 et ceux de 1978 sur la base de 1972 pour les diverses séries calculées dans cette étude et pour les séries de l'IPC officiel du Canada.

TABLEAU 11. Comparaison des indices de 1972
(1957 = 100)

	Série officielle	Série A	Série B	Laspeyres	Laspeyres en chaîne	Paasche en chaîne	Paasche	Fisher	Fisher en chaîne
Indice d'ensemble	148.2	146.3	146.0	146.3	145.4	142.4	140.8	143.5	143.9
Aliments	148.0	146.7	148.4	146.7	150.4	145.4	146.8	146.7	147.9
Aliments consommés à la maison	143.9	142.7	143.9	142.7	145.5	139.6	140.9	141.8	142.5
Repas pris à l'extérieur	N/D[1]	176.7	176.6	176.7	176.3	176.0	176.0	176.3	176.1
Ensemble sans les aliments	148.1	146.1	145.1	146.1	143.6	141.5	139.2	142.6	142.5
Habitation	150.6	151.2	150.6	151.2	148.5	146.3	143.9	147.5	147.4
Habillement	137.0	134.9	135.0	134.9	134.0	133.4	131.7	133.3	133.7
Transports	144.3	137.5	136.5	137.5	135.9	134.2	132.5	135.0	135.0
Santé et soins personnels	167.7	159.2	156.9	159.2	155.0	151.9	147.8	153.4	153.4
Loisirs, lecture et formation	156.9	152.7	151.6	152.7	150.8	145.7	138.6	145.5	148.2
Tabacs et boissons alcoolisées	140.5	142.4	141.1	142.4	140.5	140.8	139.6	141.0	140.6

[1] Chiffre non disponible sur la base 1957 = 100, cet agrégat n'ayant été inclus qu'en 1961 dans l'IPC officiel du Canada.

TABLEAU III. Comparaison des indices de 1978
(1972 = 100)

	Série officielle	Série A	Série B	Laspeyres	Laspeyres en chaîne	Paasche en chaîne	Paasche	Fisher	Fisher en chaîne
Indice d'ensemble	167.2	166.9	164.6	163.7	163.6	160.7	160.3	162.0	162.1
Aliments	193.3	196.7	197.9	197.4	201.2	193.0	195.9	196.6	197.1
Aliments consommés à la maison	194.3	197.5	199.2	198.5	203.8	194.5	197.3	197.9	199.1
Repas pris à l'extérieur	187.7	192.6	193.1	192.9	193.0	192.2	192.3	192.6	192.6
Ensemble sans les aliments	158.1	157.6	154.9	154.1	154.5	153.2	152.8	153.4	153.8
Habitation	166.3	166.1	162.5	160.5	159.9	159.7	158.9	159.7	159.8
Habillement	142.7	143.2	139.8	140.4	139.7	139.1	139.2	139.8	139.4
Transports	158.1	157.2	156.2	157.1	156.6	156.1	154.9	156.0	156.3
Santé et soins personnels	158.6	158.7	156.8	155.9	155.9	155.3	154.8	155.3	155.6
Loisirs, lecture et formation	144.2	146.5	143.5	141.8	143.2	142.5	143.7	142.7	142.8
Tabacs et boissons alcoolisées	151.4	149.6	149.5	149.5	149.5	149.5	149.5	149.5	149.5

Renvois

[1] Pour des raisons de simplicité, les indices des formules de l'annexe I sont écrites sous forme de proportions plutôt que sous forme de pourcentages, même si on emploie généralement les pourcentages quand on publie les indices.

[2] Voir la publication n° 62-546 au catalogue de Statistique Canada, *op. cit.*, section 3.5.

LINKING PRICE INDEX NUMBERS

Pour vous fournir une version dans la langue officielle de votre choix, le texte anglais est suivi du texte français (p.567) dans cette publication.

Bohdan J. Szulc
Statistics Canada
Prices Division
Central Research Section

SUMMARY

Chain indices are deemed to be less biased and more relevant than the corresponding direct fixed-basket indices. The arguments usually advanced to substantiate this superiority are not, however, very convincing. They deal with the soundness of particular fixed-basket indices that are linked into the chain, and not with the soundness of the whole chain index. The author revises these arguments and examines conditions under which a chain index is likely to perform worse than its direct counterpart.

The crucial condition is "bouncing" of relative prices in the interval from the base to the target time, which implies a specific negative intertemporal correlation between price changes. This correlation is defined in general terms in Section 4, following a decomposition of the divergence between the corresponding chain and direct indices in Section 3. Under that condition, and assuming a negative correlation between price and quantity changes, chain indices of the Laspeyres, Paasche and Sauerbeck formulae (chosen for their use in practice) tend to be no less biased than their direct counterparts, as shown in Sections 5 to 7. This is especially true when link times lie close to the inflection points of relative prices.

In the same circumstances, link times lie close to the inflection points of relative quantities, which, as explained in Section 8, implies that baskets used in linked indices lie apart from the linear combinations of those baskets that are drawn from the base and the target times. Then, however, a chain index is no more relevant than its direct counterpart, either.

Theoretical and practical conclusions from the above findings are discussed in Section 8. In actual situations, chain indices are not likely to be much more biased than the corresponding direct indices, but caution is required when linking is performed to measure price changes that exhibit cyclical variations correlated with quantity variations. This notwithstanding, there seems to be no valid workable alternative than to turn to linking fixed-basket indices when price changes have to be measured over long time intervals.

1. Formulation of the Problem

The suitability of linking price indices in time series is discussed in this paper. Linking is an operation that generates a chain index, as opposed to a direct index. A direct price index $P^D_{t/0}$ measures price change throughout a given time interval (0,t) in one leap from the base time 0 to the target (observation) time t. A chain index $P^C_{t/0}$ measures it by means of a product of indices $P_{1/0}, P_{2/2}, \ldots, P_{t/t-1}$ that quantify price changes in all sub-intervals (0,1), (1,2),..., (t-1,t) into which the given time interval has been divided:

$$P^C_{t/0} = \prod_{j=1}^{t} P_{j/j\text{-}1} \qquad (1)$$

Chain indices serve as substitutes for direct indices, particularly when the latter are judged impracticable or inadequate, as in the case of long-run index series. When the target time becomes very remote from the base time, no direct index formula can be satisfactory, in particular, no fixed-basket or equivalent formula.[1] In such circumstances, Laspeyres and Paasche indices are likely not only to diverge from each other, but to become unrealistic if not completely meaningless. Nor does an intermediate basket help, because in the case of a very large interval (0,t) the basket is necessarily drawn from a period that is remote from either 0 or t or both.

When direct comparisons are limited to sub-intervals, as in the case of a chain index, the gap between the corresponding Laspeyres and Paasche indices is likely to become narrower and an arbitrary choice between them (or a decision to take their average) is less worrisome than in a direct long-term price comparison. More important, the idea of updating baskets is very attractive, by itself, because it brings price indices closer to reality

and makes the long-term comparison look like a very natural flow of self-adjusting measurements.

A. Marshall [1887] was the first author who suggested that weights be updated periodically and indices linked, in order to make relevant price comparisons through time. An elegant theoretical rationale of the chain approach is due to F. Divisia [1925] who gave a differential definition of price and quantity indices for time sub-intervals. A chain index is then obtained by integrating these indices along the time path. Divisia pointed out that there is virtually no problem of index formula if time sub-intervals are infinitesimal or, in practical approximation, very narrow. A further incentive to link price indices came from several contemporary authors who deal with the theory of the cost-of-living index. They argue that the "true" cost-of-living indices, unfortunately not workable, can be fairly approximated by feasible Laspeyres and Paasche fixed-basket price indices, in particular, when the latter are close to each other. According to W. Diewert [1983, Section 4], "This implies that we should use the chain principle for constructing indexes rather than the fixed base principle."

Taking into account all or some of these theoretical and practical arguments, index makers often apply the linking procedure in empirical computations. The procedure, however, should not be applied indiscriminately because in certain circumstances chain indices may not measure price changes better than their direct counterparts, but rather worse. The present paper carries out an analysis of these circumstances with respect to fixed-basket price indices,[2] selected because of their frequent use in practice and because of their transparent arithmetic properties. In fact, a similar analysis is possible with respect to all price indices that can be presented as arithmetic means of indices for sub-aggregates, although the interpretation is not always easy.

2. Arithmetic Properties of Chain Indices

The basic features of a fixed-basket index can be easily derived from its formula. Most important, such an index computed for a given aggregate of commodities is always interpretable as a weighted arithmetic mean of price indices for the respective sub-aggregates. Consequently, every fixed-basket index bears all characteristics of an arithmetic mean[3] and this greatly facilitates the understanding of its behaviour. For example, no fixed-basket

index for a given aggregate can exceed (fall short of) the highest (the lowest) price indices for its sub-aggregates and one may expect that a given fixed-basket index will be particularly sensitive to the price movement of those commodities that carry large weights, i.e., that have large shares in the basket (these shares being valued in the base-time prices).

By contrast, the features of a chain price index remain rather obscure, even when all linked indices $P_{j/j-1}$ are based on a fixed-basket formula. In the latter case, the knowledge of these formulae is helpful in understanding the behaviour of particular linked indices $P_{1/0}, P_{2/1}, \ldots, P_{t/t-1}$, but does not provide any apparent clue as to the behaviour of their product, i.e., of the resulting chain index. In general, such a chain index computed for a given aggregate of commodities cannot be interpreted as an arithmetic mean of price indices for the respective sub-aggregates (unless negative value weights are allowed). For this reason chain indices may yield surprising results, as shown in the following very simple numerical example.

Example 1

Let p_0, p_1, p_2, p_3, p_4 and q_0, q_1, q_2, q_3, q_4 be prices and quantities, given for just three commodities A, B, C in five time points 0, 1, 2, 3, 4, respectively.

Commodity	Price					Quantity				
	p_0	p_1	p_2	p_3	p_4	q_0	q_1	q_2	q_3	q_4
A	3	4	5	4	3	14	12	10	12	14
B	5	5	5	5	5	10	10	10	10	10
C	7	6	5	6	7	6	8	10	8	6

Now, the Laspeyres fixed-basket price indices[4] for consecutive sub-intervals (0,1), (1,2), (2,3) and (3,4) are:

$$P^L_{1/0} = \frac{142}{134} \approx 1.06 \qquad P^L_{2/1} = \frac{150}{146} \approx 1.03$$

$$P^L_{3/2} = \frac{150}{150} = 1.00 \qquad P^L_{4/3} = \frac{142}{146} \approx 0.97$$

Their product, i.e., the chain Laspeyres index covering the entire interval (0,4) is $P^{CL}_{4/0} \approx 1.06$. It indicates a price increase of about 6%, while all prices return in the target time 4 to the same level as in the base time 0.

Example 1 may appear artificial, but the actually published series also contain some chain price indices with surprising properties. For instance, there is a case of an index being lower than indices for both its sub-components[5].

The behaviour of an obscure chain index can be elucidated, to some degree, by comparing it with a direct index of a known formula, hence of known basic features. This approach is possible, in particular, when all linked indices $P_{j/j-1}$ are computed using one index formula, say formula *, as is often done in practice. Then, the chain index $P^{C*}_{t/0}$ obtained by multiplying these indices $P^*_{1/0}, P^*_{2/1}, \ldots, P^*_{t/t-1}$ can be compared with a direct index $P^{D*}_{t/0}$ based on the same formula *. In general, the two differ from each other:

$$P^{C*}_{t/0} \neq P^{D*}_{t/0} \tag{2}$$

The proposition presented in (2) does not apply to those index formulae that call for the use of the same fixed basket in all price indices, no matter what their base time and target time.[6] The above formulae are transitive, which means precisely that the corresponding chain and direct indices are always identical. This case is not discussed in the present paper because of its trivial character and of the fact that a chain index computed without basket updating seems to contradict the spirit and purpose of linking.

The proposition presented in (2) is not sufficient to help in deciding against or in favour of the linking procedure even in those circumstances when the corresponding series of chain

and direct indices diverge quite substantially and systematically (''drift'' according to the terminology used by R. Frisch [1936]). It is possible, however, to make some progress by analysing factors and conditions that contribute to the discrepancy between the corresponding chain and direct indices.

One might be surprised to find such a limited number of papers devoted to this topic, given the extensive use of the linking procedure in practice. This notwithstanding, the methods and ideas presented here have precedents in earlier works by several authors. First of all, L. Bortkiewicz [1924] elaborated a method of decomposing the discrepancy between two price indices (Paasche and Laspeyres, as a matter of fact) that is largely used in this paper in another context. R. Frisch [1936] displayed the basic problem addressed in this paper, but developed it in detail only with respect to Sauerbeck's formula and did not pursue all consequences. V. Zarnowitz [1961] discussed those aspects of the problem that relate to the seasonality of prices and quantities, although he implicitly raised some more general issues for the first time. R. Allen [1975] was the first to extend the analysis to more than two linked indices, but only with respect to the Laspeyres formula. Finally, F. Forsyth [1978] and F. Forsyth and R. Fowler [1981] should be mentioned as the most recent contributors to the topic.

3. Factors of the Discrepancy Between the Corresponding Chain and Direct Fixed-basket Indices

The discrepancy between the corresponding chain and direct fixed-basket indices is treated in this section in a most general way, irrespective of the specific formula used.

Let $P^{C*}_{t/0}$ be a chain price index obtained by linking fixed-basket indices and let the j-th linked index be defined as follows:

$$P^{*}_{j/j-1} = \frac{\Sigma\, p_j q_{k_j}}{\Sigma\, p_{j-1} q_{k_j}} = \frac{\Sigma\, (p_j/p_{j-1})\, (p_{j-1} q_{k_j})}{\Sigma\, (p_{j-1} q_{k_j})} \qquad (3)$$

where

p_j and p_{j-1} represent prices of an individual commodity in time j and j-1 respectively,
q_{k_j} is the fixed quantity attributed to that commodity in this particular j-th linked index, drawn from the basket reference time k_j, and
Σ extends over all commodities contained in the given aggregate.

In other words:

$$P^{C*}_{t/0} = \prod_{j=1}^{t} \frac{\Sigma\,(p_j/p_{j-1})\,(p_{j-1}q_{k_j})}{\Sigma\,(p_{j-1}q_{k_j})} \tag{4}$$

Similarly, let $P^{D*}_{t/0}$ be a direct fixed-basket price index defined as follows:

$$P^{D*}_{t/0} = \frac{\Sigma\, p_t q_d}{\Sigma\, p_0 q_d} \tag{5}$$

where

p_t and p_0 are prices of an individual commodity in the target and base time, respectively,
q_d is the fixed quantity attributed to that commodity in this particular direct index, drawn from the basket reference time d, and
Σ extends over all commodities contained in the given aggregate.
The above index can be also written in another way:

$$P^{D*}_{t/0} = \prod_{j=1}^{t} \frac{\Sigma\, p_j q_d}{\Sigma\, p_{j-1} q_d} = \prod_{j=1}^{t} \frac{\Sigma\,(p_j/p_{j-1})\,(p_{j-1}q_d)}{\Sigma\,(p_{j-1}q_d)} \tag{6}$$

The ratio of the corresponding chain and direct indices, as they appear in formulae (4) and (6), can be presented as the following product:

$$P^{C^*}_{t/0} \div P^{D^*}_{t/0} = \prod_{j=1}^{t} F^*_j \tag{7}$$

with factors F^*_j being ratios of two arithmetic means of the same price relatives, though differently weighted:

$$F^*_j = \frac{\Sigma\,(p_j/p_{j-1})\,(p_{j-1}q_{k_j})}{\Sigma\,(p_{j-1}q_{k_j})} \div \frac{\Sigma\,(p_j/p_{j-1})\,(p_{j-1}q_d)}{\Sigma\,(p_{j-1}q_d)} \tag{8}$$

The proportion of alternative weights assigned to a given commodity in the above means is equal to:

$$(p_{j-1}q_{k_j}) \div (p_{j-1}q_d) = q_{k_j}/q_d$$

i.e., to the quantity relative for this commodity between the two fixed-basket reference times k_j and d, respectively. To simplify the algebra, let these quantity relatives be designated as y's:

$$y = q_{k_j}/q_d \tag{9}$$

and the averaged price relatives as x's:

$$x = p_j/p_{j-1} \tag{10}$$

Now, each of the factors F^*_j is open to the following decomposition, similar to that suggested by Bortkiewicz [1924]:[7]

$$F^*_j = 1 + r_{xy} \cdot V_x \cdot V_y \qquad \text{for } j = 1, 2, \ldots, t \tag{11}$$

where

r_{xy} is the coefficient of linear correlation between the respective price relatives x and quantity relatives y,

V_x is the coefficient of variation of price relatives x, and

V_y is the coefficient of variation of quantity relatives y.

It follows from (7) that the chain index $P^{C*}_{t/0}$ is higher (lower) than the corresponding direct index $P^{D*}_{t/0}$ if factors F^*_j are predominantly above (below) unity which, according to (11), depends solely upon the sign of the correlation coefficients r_{xy}.[8]

4. Correlation Between Current Price Changes and Cumulated Quantity Changes

The question of the direction in which the chain index $P^{C*}_{t/0}$ diverges from its direct counterpart $P^{D*}_{t/0}$ has been reduced to the problem of signs of the correlation coefficients r_{xy} within the universe of commodities contained in the given aggregate. Each coefficient expresses the correlation between price relatives $x = p_j/p_{j-1}$ and quantity relatives $y = q_{k_j}/q_d$. Price relatives x measure **current** price changes that occur in consecutive subintervals (j-1,j), immediately preceding any selected link time j. Quantity relatives y, instead, measure quantity changes that **cumulate** during time intervals (d,k_j) or (k_j,d), i.e., between the reference time of the basket used in the direct index $P^{D*}_{t/0}$ (a constant) and the reference times of baskets used in linked indices $P^*_{j/j-1}$ (in general, a variable). Since x and y do not refer to the same time interval, it is not easy to conjecture what signs would prevail in coefficients r_{xy} under given circumstances. Some authors, indeed, seem to hesitate in this respect (Allen [1975, pp. 197-199]), while some others tend to take a rather one-sided position (Frisch [1936, p.9] commenting on Laspeyres and Paasche indices).

The above difficulties can be overcome in part by analysing other, more intelligible coefficients of correlation. The suggested approach takes advantage of a specific relationship that holds for any triad of coefficients of correlation r_{xy}, r_{xz} and r_{yz} computed for the same population. In fact, the limits of r_{xy} can be written as the following function of r_{xz} and r_{yz}:[9]

$$r_{xz} \cdot r_{yz} - [(1-r^2_{xz})(1-r^2_{yz})]^{1/2} \leq r_{xy} \leq r_{xz} \cdot r_{yz} + [(1-r^2_{xz})(1-r^2_{yz})]^{1/2} \quad (12)$$

The double inequality (12) is not a very efficient way of inferring the value of r_{xy}, especially when both correlations x to z and y to z are weak. On the other hand, if either of the two correlations becomes stronger, the limits of r_{xy} tend to narrow and, in particular, the sign of the r_{xy} tends to be determined by the combination of signs of r_{xz} and r_{yz}. When both r_{xz} and r_{yz} are of the same sign, r_{xy} tends to be positive rather than negative, while the contrary happens when r_{xz} and r_{yz} are of opposite signs. Consequently, the knowledge of the signs of r_{xz} and r_{yz} could be helpful in conjecturing about the sign of r_{xy}.

Now, define the variable z as follows:

$$z = p_{k_j}/p_d \quad (13)$$

The coefficient r_{yz}, with y defined by (9), measures the correlation between quantity changes $y = q_{k_j}/q_d$ and price changes $z = p_{k_j}/p_d$ for particular commodities, that occurred in the same time interval (d, k_j). Due to price-induced substitution, the coefficients r_{yz} tend to be negative for consumer goods, whatever this interval. According to (12), this means that the signs of coefficients r_{xy}, under consideration, depend to a large degree upon the signs of coefficients r_{xz}. The latter measure the intertemporal correlation between the current price changes $x = p_j/p_{j-1}$ and price changes $z = p_{k_j}/p_d$ that cumulate from time d (fixed) to time k_j (moving). If r_{xz}'s tend to be negative for $j = 1,2,\ldots,t$, then r_{xy}'s tend to be positive and the chain index $P^{C*}_{t/0}$ tends to be higher than the corresponding direct index $P^{D*}_{t/0}$. The opposite is true if r_{xz}'s tend to be positive.

The correlation x to z is an intertemporal correlation between price changes.[10] The sign of the coefficient r_{xz} depends, therefore, on the pattern of price changes during periods under consideration. It also depends on the choice of time points j, k_j and d for the purpose of computing the compared chain and direct indices. Time points $j = 1,2,\ldots,t$ are the limits of sub-intervals (0,1), (1,2),..., (t-1,t) into which time interval (0,t) has been divided. They are link points at which the consecutive indices $P^*_{j/j-1}$ are linked to those following. Time points k_j (several, as a rule) and d (unique) are time references of fixed baskets

used in the linked indices $P^*_{j/j-1}$ and in the direct index $P^{D*}_{t/0}$, respectively. They depend upon the index formula, hence the intertemporal correlation between price changes must be discussed in the context of the formula chosen. This will be done in the following two sections with respect to the Laspeyres and Paasche formulae, selected as examples because of their importance in index number theory and practice.

The sign of r_{xz} influences the sign of r_{xy} through the medium of r_{yz} and its impact is moderated by the weakness of the correlation y to z, if this is the case. As stated above, r_{yz}'s are rather negative for consumer goods, but they are rarely very strong. In fact, price-induced substitution is not the sole force causing changes in consumption (see P. Généreux [1983]). There are, however, some linking practices that create mechanisms imitating a thorough and extreme price-induced substitution. Then, the sign opposite to that of r_{xz} is definitely transmitted to r_{xy}. This is the case of linking indices with the Sauerbeck formula, discussed after that of the Laspeyres and Paasche formulae.

5. Intertemporal Correlation Between Price Changes and the Laspeyres Chain Index

In the case of Laspeyres indices, the compared formulae are:

$$P^{D*}_{t/0} = P^{DL}_{t/0} = \frac{\Sigma\, p_t q_0}{\Sigma\, p_0 q_0} \qquad \text{and}$$

$$P^{C*}_{t/0} = P^{CL}_{t/0} = \prod_{j=1}^{t} \frac{\Sigma\, p_j\ q_{j-1}}{\Sigma\, p_{j-1} q_{j-1}}, \qquad \text{which implies}$$

$$q_{k_j} = q_{j-1} \text{ and } q_d = q_0 \text{, as well as}$$

$$x = p_j/p_{j-1},\ y = q_{j-1}/q_0 \text{ and } z = p_{j-1}/p_0$$

Note that for j = 1 there is $q_{k_j} = q_d$, hence, according to (8), $F^L_1 = 1$. For all remaining j's, as stated in the previous section, the F^L_j's tend to depend on the sign of the corresponding r_{xz}. In the case of Laspeyres indices, the r_{xz} corresponding to a given F^L_j measures the correlation between current price changes p_j/p_{j-1} and price changes p_{j-1}/p_0 accumulated from the base time 0. In other words, this correlation refers to price changes

in the following pair of time intervals: (j-1,j) and (0,j-1). For the consecutive j = 2,3,...,t, these pairs are:

j	Intervals corresponding to		
	p_{j-1}/p_0	and	p_j/p_{j-1}
2	(0,1)	and	(1,2)
3	(0,2)	and	(2,3)
		.	
		.	
		.	
t-1	(0,t-2)	and	(t-2,t-1)
t	(0,t-1)	and	(t-1, t)

Consider the pattern of relative price changes from the base time 0 to the target time t. If the same tendency of relative price changes persisted during this entire period,[11] i.e., if the rank of commodities by their relative price was maintained during this entire period, all r_{xz}'s for j = 2,3,...,t would be positive. Assuming negative r_{yz}, the above implies that the chain Laspeyres index would tend to be lower than its direct counterpart. This tendency becomes less strong with less frequent linking.

On the other hand, if between the base time 0 and the target time t relative prices tend to "bounce", some r_{xz}'s would be negative and the chain Laspeyres index would tend to be higher than its direct counterpart. If only two indices are linked, the impact of "bouncing" is the stronger the closer time 0 and time t lie to the peaks, and link points j to the troughs of relative price movements (or vice versa). If more links are added within the same cycle of relative price movements, the impact will be, generally, less pronounced.

When the relative price movement exhibits repetitive cycles and linking time points are selected at (or close to) the consecutive peaks and troughs, the chain Laspeyres index may drift from its direct counterpart indefinitely. The synchronization of linking phases with those of relative price changes has an effect on the chain index analogous to the synchronization of wind phases with inborn periodicity of movements of the Tacoma Narrows suspension bridge, which collapsed in 1940.

Below is shown a numerical example of a persistent tendency of relative prices, combined with a negative correlation between price and quantity changes:

Commodity	Price					Quantity				
	p_0	p_1	p_2	p_3	p_4	q_0	q_1	q_2	q_3	q_4
A	3	4	5	6	7	14	12	10	8	6
B	5	5	5	5	5	10	10	10	10	10
C	7	6	5	4	3	6	8	10	12	14

The Laspeyres fixed-basket price indices for the consecutive time sub-intervals are:

$$P^L_{1/0} = \frac{142}{134} \approx 1.06 \qquad P^L_{2/1} = \frac{150}{146} \approx 1.03$$

$$P^L_{3/2} = \frac{150}{150} = 1.00 \qquad P^L_{4/3} = \frac{142}{146} \approx 0.97$$

Their product, i.e., the chain Laspeyres index for the entire time interval (0,4) is $P^{CL}_{4/0} \approx 1.06$, lower than the corresponding direct Laspeyres index $P^{DL}_{4/0} = \frac{166}{134} \approx 1.24$. The difference becomes smaller if just one linking (instead of the previous three) was performed, e.g., at the time point 2. In that latter case, the linking Laspeyres indices are: $P^L_{2/0} = \frac{150}{134} \approx 1.12$ and $P^L_{4/2} = \frac{150}{150} = 1.00$, which leads to the chain Laspeyres index $P^{CL}_{4/0} \approx 1.12$.

Example 1 (Section 2) is an illustration of "bouncing" relative prices (their tendency being reversed in time 2) with a negative correlation between price and quantity changes. If indices are linked only once, just at the time point 2, then the linked Laspeyres indices are: $P^L_{2/0} = \frac{150}{134} \approx 1.12$ and $P^L_{4/2} = \frac{150}{150} = 1.00$, which leads to the chain Laspeyres index $P^{CL}_{4/0} \approx 1.12$, higher than the corresponding direct Laspeyres index $P^{DL}_{4/0} = 1.00$. The difference becomes smaller, even though it remains of the same sign, if linking is performed thrice, at time points 1, 2 and 3. In this case, the chain Laspeyres index is 1.06, as shown

in Section 2. Note that the link time 1 falls within the interval (0,2), during which the same tendency of relative prices prevailed. The link time 3 occupies a similar position within the interval (2,4). Finally, with just one linking but at the time point 1, the linked Laspeyres indices are $P^L_{1/0} = \frac{142}{134} \approx 1.06$ and $P^L_{4/1} = \frac{142}{146} \approx 0.97$ and the chain index is $P^{CL}_{4/0} \approx 1.03$, less than with linking at the reversal time 2.

Example 3 (Section 7) may serve as an illustration of repetitive cycles in the relative price movement and of its impact on chain indices.

6. Intertemporal Correlation Between Price Changes and the Paasche Chain Index

In the case of Paasche indices, the compared formulae are:

$$P^{D*}_{t/0} = P^{DP}_{t/0} = \frac{\Sigma\, p_t q_t}{\Sigma\, p_0 q_t} \text{ and}$$

$$P^{C*}_{t/0} = P^{CP}_{t/0} = \prod_{j=1}^{t} \frac{\Sigma\, p_j q_j}{\Sigma\, p_{j-1} q_j}, \text{ which implies}$$

$$q_{k_j} = q_j \text{ and } q_d = q_t, \text{ as well as}$$

$$x = p_j/p_{j-1},\ y = q_j/q_t \text{ and } z = p_j/p_t$$

Note that for $j=t$ there is $q_{k_j} = q_d$, hence, according to (8), $F^P_t = 1$. For all remaining j's, as stated in the previous section, the F^P_j's tend to depend on the sign of the corresponding r_{xz} or on the sign opposite to the corresponding $r_{x\,1/z}$. The latter is the coefficient that measures the correlation between current price changes p_j/p_{j-1} and price change p_t/p_j cumulated until the target time t. In other words, this correlation refers to price changes in the following pair of time intervals: (j-1,j) and (j,t). For the consecutive $j=1,2,\ldots,t-1$, these pairs are:

j	Intervals corresponding to		
	p_j/p_{j-1}	and	p_t/p_j
1	(0,1)	and	(1,t)
2	(1,2)	and	(2,t)
.		.	
.		.	
.		.	
t-2	(t-3,t-2)	and	(t-2,t)
t-1	(t-2,t-1)	and	(t-1,t)

Consider the pattern of relative price changes from the base time 0 to the target time t as in the case of Laspeyres indices. If the same tendency of relative price changes persisted during this entire period, all $r_{x\ 1/z}$'s for $j = 1,2,\ldots,t\text{-}1$ would be positive and, consequently, all r_{xz}'s would be negative (since neither x nor z can normally take a negative value). Assuming negative r_{yz}'s, the above implies that under those circumstances the chain Paasche index would tend to be higher than its direct counterpart.

If relative prices tend to "bounce" between the base time 0 and the target time t, some $r_{x\ 1/z}$'s would be negative, the corresponding r_{xz}'s positive and the chain Paasche index would tend to be lower than its direct counterpart. The impact of "bouncing" is conditioned by the position of 0, j and t, in a similar way as in the case of Laspeyres indices, discussed in Section 5.

In the numerical Example 2 (Section 5), relative price changes exhibit a persistent tendency and are negatively correlated with quantity changes. The direct Paasche index is $P^{DP}_{4/0} = \frac{134}{166} \approx 0.81$. The chain Paasche index with linking at three time points 1, 2 and 3 is: $P^{CP}_{4/0} \approx 1.03 \times 1.00 \times 0.97 \times 0.93 \approx 0.94$, more than the direct one. The chain Paasche index with a single linking at the time point 2 is: $P^{CP}_{4/0} \approx 1.00 \times 0.89 \approx 0.89$ and at the time point 1: $P^{CP}_{4/0} \approx 1.03 \times 0.85 \approx 0.87$, also more than the direct Paasche index, although the difference is now less pronounced.

In the numerical Example 1 (Section 2) relative prices "bounce" at the time point 2 and are also negatively correlated with quantity changes. The direct Paasche index is $P^{DP}_{4/0} = \frac{134}{134} = 1.00$. The chain Paasche index with linking at the time points 1, 2 and 3 is: $P^{CP}_{4/0} \approx 1.03 \times 1.00 \times 0.97 \times 0.94 \approx 0.94$, less than the direct one. The chain Paasche index with a single linking at the time point 2 is: $P^{CP}_{4/0} \approx 1.00 \times 0.89 \approx 0.89$ and at the time point 1 is: $P^{CP}_{4/0} \approx 1.03 \times 0.94 \approx 0.97$, also less than the direct index. The largest difference is shown when linking is performed at the time point 2 (the reversal of the relative price change tendency).

7. Intertemporal Correlation Between Price Changes and the Sauerbeck Chain Index

Now, take as an example a simple (unweighted) arithmetic mean of price relatives, referred to as the Sauerbeck index. The formula is sometimes used in practice, though not at higher levels of aggregation but, rather, to combine individual price data (collected, say, every month) into the lowest-level month-to-month price indices. When there is a change in items selected to represent the price movement of the given lowest-level aggregate, these indices can be linked. Moreover, in order to have a uniform algorithm of index computation, the linking procedure can be extended to all lowest-level month-to-month Sauerbeck indices, whether there has been a change in representative items or not. The chain Sauerbeck index, however, has very peculiar properties.

Let us compare the chain and direct Sauerbeck indices:

$$P^{D*}_{t/0} = P^{DS}_{t/0} = \frac{\Sigma\,(p_t/p_0)}{N} \quad \text{and}$$

$$P^{C*}_{t/0} = P^{CS}_{t/0} = \prod_{j=1}^{t} P^{S}_{j/j-1} = \prod_{j=1}^{t} \frac{\Sigma\,(p_j/p_{j-1})}{N},$$

where N is the number of items for which price relatives are averaged in each index. It is possible to transform the above direct and linked indices into implicit fixed-basket indices, as follows:

$$P^{DS}_{t/0} = \frac{\Sigma\,(p_t/p_0)}{N} = \frac{\Sigma\,(p_t/p_0)(p_0\frac{1}{p_0})}{\Sigma\,(p_0\frac{1}{p_0})} \text{ and}$$

$$P^{S}_{j/j-1} = \frac{\Sigma\,(p_j/p_{j-1})}{N} = \frac{\Sigma\,(p_j/p_{j-1})(p_{j-1}\frac{1}{p_{j-1}})}{\Sigma\,(p_{j-1}\frac{1}{p_{j-1}})}$$

We conclude that the direct and linked Sauerbeck indices are equivalent to fixed-basket price indices, in which implicit fixed quantities are: $q_d = \frac{1}{p_0}$ and $q_{k_j} = \frac{1}{p_{j-1}}$. This means that, in every time interval (0, j-1), the variation of implicit fixed quantities is the reciprocal of the variation of prices. In other words, Sauerbeck indices correspond to fixed-basket price indices that are computed under the assumption of an extremely strong price-induced substitution, working throughout all items.

In these circumstances, all r_{yz}'s are equal to minus one and every r_{xy} is unequivocally determined by r_{xz} (with the opposite sign) or, which is easier to follow, by $r_{x1/z}$ (with the same sign). The latter coefficient measures the correlation between current price changes p_j/p_{j-1} and cumulated price changes p_{j-1}/p_0. They refer to time intervals (j-1, j) and (0, j-1), the same as in the case of Laspeyres indices.

If the same tendency of relative price changes persisted during the entire period from the base time 0 to the target time t, all r_{xy}'s would be negative. This implies that the chain Sauerbeck index would be lower than its direct counterpart. The opposite holds true when relative prices tend to "bounce". From this point of view, the relationship between the corresponding chain and direct Sauerbeck indices is analogous to that of Laspeyres indices. With the same pattern of relative price changes, however, their impact on a Sauerbeck chain index is generally much stronger. With repetitive "bouncing", and the linking cycle synchronized with that of relative price movement, the results may be perplexing. This is illustrated by the following numerical example, with just two items for which price relatives are calculated.

Item	Price					Price relative			
	p_0	p_1	p_2	p_3	p_4	p_1/p_0	p_2/p_1	p_3/p_2	p_4/p_3
A	1	2	1	2	1	2.0	0.5	2.0	0.5
B	2	1	2	1	2	0.5	2.0	0.5	2.0
Simple Mean						1.25	1.25	1.25	1.25

The chain Sauerbeck index is:

$$P^{CS}_{0/4} = \prod_{j=1}^{4} \frac{\Sigma\,(p_j/p_{j-1})}{2} = 1.25^4 \approx 2.44,$$

which indicates a price increase of 144% with no actual price change at all from time 0 to time 4. Note that if linked indices were defined explicitly as fixed-basket indices with quantity "one" attached to each item (or as ratios of simple arithmetic means of prices, which is equivalent to the above), the chain index would be equal to unity:

$$\prod_{j=1}^{4} \frac{\Sigma\, p_j \cdot 1}{\Sigma\, p_{j-1} \cdot 1} = \left(\frac{3}{3}\right)^4 = 1$$

8. Suitability of Chain Indices

The suitability of chain indices cannot be assessed without examining whether, and on what occasions, they are likely to exhibit those features that are the main objective of the linking procedure. The divergence, if any, between chain and direct indices is not a generally valid basis for such an assessment. When the linking procedure is employed in practice, direct indices are usually deemed to be doubtful measures of price changes, at best; if they were not, linking would be unnecessary. Accordingly, the fact that a given chain index diverges from the corresponding direct index is not, per se, symptomatic of its "wrongness" or "correctness". In some cases, however, it is not only possible but advisable to judge the inadequacy

of a given chain index by the direction in which it diverges from the direct one. This judgment, to be well-grounded, should pertain to the assessed bias of the corresponding direct index.

The term "bias" denotes a deviation between the result yielded by a given index and that expected of the theoretically correct index. In the case of consumer prices, cost-of-living indices constitute such benchmarks. Because of the substitution effect, the following relationships will prevail:

- Laspeyres-type cost-of-living indices tend to be lower than the corresponding Laspeyres fixed-basket price indices and, in that sense, the latter are considered upward-biased,
- Paasche-type cost-of-living indices tend to be higher than the corresponding Paasche fixed-basket price indices and, in that sense, the latter are considered downward-biased.

The longer the time interval, the more substantial are the changes in price structure and consumption pattern that are likely to occur, with more potential for the substitution effects. Consequently, the respective upward and downward biases of Laspeyres and Paasche fixed-basket price indices tend also to grow as the target time becomes distant from the base time.

When the linking procedure is used, fixed-basket indices are calculated for relatively narrow sub-intervals of time, hence their biases are likely to remain moderate whatever the distance between the base and the target time of the chain index. On that premise, it is quite commonly alleged that, in the case of elongated time intervals, chain indices are less prone to growing bias than their direct counterparts. This, in turn, is one of the main reasons for employing chain indices in practice.

There is, however, a notable pitfall in logic in the above allegation. The global bias of the entire chain index can be large even if the biases of particular linked indices are moderate. According to Sections 5 and 6 of this paper, a chain of Laspeyres fixed-basket indices can be higher than the corresponding direct Laspeyres price index and a chain of Paasche fixed-basket indices can be lower than the corresponding direct Paasche price index. In other words, chain indices can be more biased than their direct counterparts, which is both contrary to popular belief and to the purpose of linking. Linking, therefore, should be avoided in a situation when such undesirable results tend to materialize.

As proven in previous sections, this is likely to happen when:

- there is a negative correlation between price and quantity changes, and
- relative prices ''bounce'' in the interval from the base to the target time.

In practice, the second condition requires more attention, because in most real life situations there is some negative correlation between price and quantity changes, anyway, which is due to price-induced substitution of commodities by consumers. For the same reason, the second condition is generally equivalent to the following:

- relative quantities ''bounce'' in the interval from the base to the target time.[12]

The bias of chain indices grows with stronger correlation between price and quantity changes. Hence, special attention should be paid to those situations when such a correlation is likely to occur and in particular, when price-induced substitution is known to be highly pronounced. Seasonal price and quantity changes are a typical example, while no correlation may occur if price or quantity changes happen just randomly. The Sauerbeck formula, which triggers this correlation automatically, albeit in a concealed way, should be definitely avoided in the linking procedure.[13]

The bias of chain indices is stronger the closer the link points lie to the peaks of relative price (and quantity) movement, and the closer the base and the target time lie to the troughs of this movement, or vice versa. Such a synchronization of linking periodicity with that of relative prices (or quantities) should be especially avoided. If this advice is ignored and similar synchronization maintained over continuing cycles, the chain index bias could grow indefinitely. On the other hand, if more links are added within a particular cycle of the relative price (or quantity) movement, the bias would be smaller.[14]

In practice, chain indices are unlikely to exhibit much larger bias than the corresponding direct indices. This is particularly true with respect to aggregates composed of numerous commodities, for the following reasons:

- relative price (or quantity) changes are usually of a mixed character, some maintain the same tendency through time and some ''bounce'';
- when relative prices (quantities) ''bounce'', it rarely happens that one cycle is common to all commodities, hence it is nearly impossible to have a full synchronization with the linking cycle;

– the linking procedure is often employed with different formulae and periodicities at different levels of commodity aggregation; this makes the final result perhaps less transparent but also less subject to extreme behaviour than in the case of a uniform linking;

– particular linked indices are mostly based on fixed baskets drawn from time points that precede the linking points; consequently, the coefficients of intertemporal correlation measure the relationship between price (or quantity) changes in disjointed time intervals, the relationship which is unlikely to be strong;

– price-induced substitution is not a phenomenon that would show for all commodities at all times; hence, the negative correlation between price and quantity changes is normally weak.

* * *

Another main reason for employing chain indices in practice is their assumed conceptual advantage over direct indices, particularly in the case of long-run index series. In this case, sizeable changes in the distribution of consumption, production, trade, etc. (in quantitative terms) tend to appear from the base to the target time. With no basket being even approximately relevant to both these times, direct fixed-basket indices are deemed to be unrealistic, if not completely meaningless, whatever formula is applied. In such circumstances, the linking procedure seems to offer a very valid solution to the problem, by substituting a chain of rather realistic measurements of price changes for a single, unrealistic one.

The above quite popular opinion, however, conceals another logical pitfall. The linking procedure is conceived to provide a measure of price changes over the entire interval, from the base time to the target time. Price indices for particular sub-intervals of this interval, though necessary steps and by-products of the linking procedure, are not its main goal. Consequently, conceptual merits of the linking procedure should be gauged by the soundness of the whole chain index and not by that of particular linked indices used in its computation.

From this point of view, chain indices may be considered superior to their direct counterparts when they provide a smooth passage between the base and the target time, rather than a detour. It is logical and advantageous to link price indices at those time points, if any, that correspond to quantitative distributions of consumption (production, trade, etc.) lying on a path of gradual changes from the base to the target time. Speaking more strictly, these distributions should be as close as possible to linear combinations of the distributions that appear in the base and in the target time. Conversely, it is illogical to link price indices at such points

of time that correspond to quantitative distributions known to lie well outside of these paths of gradual changes.[15] This way of linking would be analogous to a comparison of distinct situations by intermedium of still more dissimilar ones or, in the extreme case, to a comparison of identical situations by intermedium of quite distinct ones.[16]

Changes in the distribution of consumption (production, trade, etc.) by commodities that result from structural changes in the economy and the population, from altered technology and habits, as well as from other long-term mutations, are normally gradual. Bearing in mind these changes, it would be beneficial to update fixed baskets and to link price indices as frequently as possible. Quantitative distributions, however, are subject to periodic, seasonal and random variations and this would call for caution with respect to the choice of fixed baskets for linked indices and of link points, themselves. In fact, and notwithstanding Divisia's advice, there is usually some minimum length of sub-interval of time that should be observed for linking purposes.

This length may vary from one area of price index making to another. For example, in the case of livestock price indices, one could try to avoid linking within hog and cattle price cycles. In the case of consumer prices, a year seems to be a theoretical optimum, although less frequent linking may well be advisable for practical reasons.[17] A more frequent linking, would build seasonal and random variations in relative quantities (and accompanying variations in relative prices) into the construction of chain indices, even though they are designed to measure long-term price movement. When random variations are likely to be substantial, it might be advantageous to base fixed baskets for linked indices on average data derived from several consecutive years. Moving averages could be applied to allow for more frequent updates, if otherwise useful, without the disturbing impact of random variations.

* * *

It is interesting, indeed important, to realize that a path of gradual changes in quantitative distributions is also the locus where no "bouncing" of relative quantities occurs.[18] Furthermore, if price and quantity changes are correlated, relative prices do not tend to "bounce" in this locus, either. In other words, both the discussions on the bias of chain indices and on their conceptual virtues lead to the same warning: caution is required when the linking procedure is employed to measure price changes that exhibit cyclical variations correlated with those of quantity variations.

This notwithstanding, there likely is no practical alternative other than linking when a long-run index series has to be established. All disadvantages of chain indices, including their somewhat peculiar arithmetic properties and the necessity of a cautious application in specific cases, seem to be minor compared to those of direct indices calculated over long time intervals.[19]

Footnotes

[1] Direct constant-standard-of-living price indices (i.e., direct cost-of-living indices) have similar weaknesses, although not that manifest. In fact, observed satisfaction levels and tastes do vary in time and there is always a Laspeyres-type and a Paasche-type cost-of-living index.

[2] *Mutatis mutandis*, the conclusions can be extended to fixed-prices volume (quantity) indices, and also to the interspatial comparisons.

[3] Note that in actual computation of the indices referred to as fixed-basket indices, a formula equivalent to the fixed-basket approach applies usually only from a certain commodity aggregation level up (see, for example, *The Consumer Price Index Reference Paper*, Statistics Canada, Cat. No. 62-553, p.27).

[4] In this paper, the term Laspeyres index is used only in the strict sense, i.e., only when fixed quantities are drawn from the base time of the index.

[5] In the Canadian Consumer Price Index series the following chain indices are published for March 1978:

Goods	171.1
Services	171.4
Goods and Services	170.8

i.e., the index for the aggregate is lower than indices for both components (see *Consumer Prices and Price Indexes*, Statistics Canada, Cat. No. 62-010, October-December 1978, p.40)

[6] The formulae in question are nowadays quite often referred to as "Laspeyres formulae", notwithstanding the original strict terminology (see Footnote 4). This lax appellation is rather regrettable. In fact, it invalidates the following well-known and very useful statements:
- the product of a Laspeyres (Paasche) price index and the corresponding Paasche (Laspeyres) volume index is equal to the ratio of current values expenditures),
- a Fisher price (volume) index is a geometric mean of the corresponding Laspeyres and Paasche price (volume) indices.

Moreover, a laxly defined series of "Laspeyres indices":

$$\Sigma\, p_t q_c / \Sigma\, p_0 q_c$$

contains necessarily a Paasche index for $t = c$, which is another indication that this lax terminology is indeed unfortunate.

[7] The decomposition is shown in Appendix 1.

[8] A case of either V_x or V_y being null is not worth particular attention. This would be a trivial case when either there is no rélative price changes in consecutive sub-intervals (j-1,j) or the proportions of commodities in all baskets used for computing linked indices are the same as in the basket used for calculating the direct index.

[9] This double inequality can be derived directly from the formula of the coefficient of partial correlation $r_{xy.z}$, knowing that $r^2_{xy.z} \geq 1$.

[10] An alternative definition of the variable z is also possible. Instead of (13), one may put:

$$z = q_j/q_{j-1}$$

In this case, the correlation y to z is an intertemporal correlation between quantity changes, while the correlation x to z is a (negative) correlation between price and quantity changes for the same time interval. Both definitions lead to analogous conclusions.

[11] If there is no relative price change, then neither is there an index problem.

[12] If relative quantities are correlated with relative prices which "bounce", they will also tend to "bounce". The modified condition can be derived following Footnote 10, as well.

[13] As previously stated, the Sauerbeck formula is sometimes used to combine individual price data, collected for particular commodities in particular outlets. Some "bouncing" of these relative prices is very likely to occur. If certain prices increase less than others in a given month or quarter, say because of special discounts offered to customers, they will tend to increase more than others in the next month or quarter.

[14] The above conclusions could be drawn directly from formulae discussed in Sections 5 and 6. A simulation of the linking procedure was performed in the Central Research Section of Prices Division, Statistics Canada. The simulation was based on price and quantity data with various seasonal patterns but without changes from year to year. Chain indices were computed with one, two, three, four, six and 12 links a year, with varying link months, where applicable, and varying base months (the target time was set always 12 months after the base month). Computer printouts and graphs confirmed all theoretical conclusions.

[15] When there is a need for both the long-term and the short-term price change measurements, the requirements with respect to the two price index series may well be inconsistent. For example, a decent price comparison between any consecutive months would require that the respective basket be as close as possible to the quantitative distributions in these months. On the other hand, a decent price comparison over a year would suffer from linking in particular months, with seasonal baskets used in the linked indices. In such cases, a decision has to be made on which price comparison is more important and requirements chosen accordingly. If both goals are equally important, a series of month-to-month indices may be produced independently of the series of year-to-year indices. Unfortunately, most index formulae are not transitive and a yearly index would provide, in general, a different result than the product of monthly indices.

[16] Indirect price comparisons derived from a Paasche price index series are a good example of such comparisons. It can be proven (see Appendix 2) that they are equivalent to a chain index in which two consecutive years (months, etc.) are compared by intermedium of the base year (month...) of the initial Paasche price index series. This might be the reason for which such indirect comparisons are so strongly abhorred, indeed often denied the name of price indices. Otherwise it would be hard to understand why the result of dividing two price indices that relate to different baskets is less a price index than the result of multiplying such indices (i.e., than a regular chain index), neither of them being a fixed-basket price index.

[17] Linking of consumer price indices every year would require substantial resources and could provide only marginally improved results compared to a chain index with links, say, every three years (see P. Généreux [1983]).

[18] See Appendix 3.

[19] In some cases, however, arithmetic properties may be decisive. For example, it is very advantageous to have values in real terms that are additive both by commodities and through time. If this is the major concern, chain price indices are not applicable for deflation purposes.

Appendix 1

Bortkiewicz [1924] has shown that a relative divergence between the corresponding Paasche and Laspeyres price indices can be transformed as follows:

$$\left[\frac{\Sigma\, p_t q_t}{\Sigma\, p_0 q_t} - \frac{\Sigma\, p_t q_0}{\Sigma\, p_0 q_0}\right] \div \frac{\Sigma\, p_t q_0}{\Sigma\, p_0 q_0}$$

$$= \left[\frac{\Sigma(p_t/p_0)(p_0 q_t)}{\Sigma(p_0 q_t)} - \frac{\Sigma(p_t/p_0(p_0 q_0)}{\Sigma(p_0 q_0)}\right] \div \frac{\Sigma(p_t/p_0)(p_0 q_0)}{\Sigma(p_0 q_0)}$$

$= r_{xy}\, V_x\, V_y$, where:

$x = p_t/p_0$ is the price relative for a particular commodity,

$y = (p_0 q_t) \div (p_0 q_0) = q_t/q_0$ is the quantity relative for this commodity,

Σ indicates the summation over all commodities (the same in both indices),

V_x is the coefficient of variation of price relatives,

V_y is the coefficient of variation of quantity relatives, and

r_{xy} is the coefficient of linear correlation between these price and quantity relatives.

The original transformation by Bortkiewicz is a special case of a more general decomposition of a relative divergence between any two arithmetic means of the same variable, say x, but weighted with alternative sets of weights, say m and n. Let the ratio of the alternative weights m/n be y.

$$\left[\frac{\Sigma xm}{\Sigma m} - \frac{\Sigma xn}{\Sigma n}\right] \div \frac{\Sigma xn}{\Sigma n} = \left[\frac{\Sigma\, x \frac{m}{n} n}{\Sigma \frac{m}{n} n} - \frac{\Sigma xn}{\Sigma n}\right] \div \frac{\Sigma xn}{\Sigma n} =$$

$$\left[\frac{\Sigma xyn}{\Sigma n} - \frac{\Sigma xn}{\Sigma n} \cdot \frac{\Sigma yn}{\Sigma n}\right] \div \left[\frac{\Sigma xn}{\Sigma} \cdot \frac{\Sigma yn}{\Sigma n}\right] = r_{xy} \cdot \frac{\sigma_x}{\bar{x}} \cdot \frac{\sigma_y}{\bar{y}} = r_{xy}\, V_x\, V_y \,.$$

It is assumed that all statistics are weighted with weights n, which correspond to (p_0q_0) in the original transformation by Bortkiewicz.

The above decomposition applies to all pairs of composite price indices that can be presented as arithmetic means of the same set of price relatives though differently weighted, as in the case of expressions $F^*_{.j}$, defined by formula (8) in the paper. The decomposition provides an interesting way of interpreting the relationship between composite price indices and, in particular, fixed-basket price indices. It has to be noted, however, that this interpretation is correct only when price and quantity relatives appearing in the coefficients r_{xy}, V_x and V_y are pertinent to that lowest level of commodity aggregation at which different weights are applied in the calculation of the compared composite price indices. By definition, the same price relatives are applied at this level of commodity aggregation in the calculation of both composite price indices.

Mutatis mutandis, the decomposition in question is also applicable to all pairs of composite volume indices that can be presented as arithmetic means of the same set of quantity relatives, albeit differently weighted (hence, to all pairs of composite volume indices based on alternative sets of fixed prices).

Appendix 2

$$\frac{\Sigma\, p_t q_t}{\Sigma\, p_0 q_t} \div \frac{\Sigma\, p_{t-1}\, q_{t-1}}{\Sigma\, p_0 q_{t-1}} = \frac{\Sigma\, p_t q_t}{\Sigma\, p_0 q_t} \cdot \frac{\Sigma\, p_0 q_{t-1}}{\Sigma\, p_{t-1} q_{t-1}}$$

Appendix 3

Let q_{0i}, q_{1i}, q_{2i} be the quantities of commodity i consumed (produced, traded etc.) in time 0, 1, 2, respectively.

Let q_{0j}, q_{1j}, q_{2j} be the analogous quantities of commodity j.

If time point 1 corresponds to a quantitative distribution lying on a path of gradual changes from time 0 to time 2, then there is:

$$\frac{q_{1j}}{q_{1i}} = \alpha \frac{q_{0j}}{q_{0i}} + (1\text{-}\alpha).\frac{q_{2j}}{q_{2i}} \quad , \qquad \text{with } 0 < \alpha < 1$$

which can be transformed into:

$$1 = \alpha \div \left[\frac{q_{1j}}{q_{0j}} \div \frac{q_{1i}}{q_{01}}\right] + (1\text{-}\alpha) \cdot \left[\frac{q_{2j}}{q_{1j}} - \frac{q_{2i}}{q_{1i}}\right]$$

and written in a simplified form:

$$1 = \alpha \div z + (1\text{-}\alpha) \cdot x$$

where:

$z = \dfrac{q_{1j}}{q_{0j}} \div \dfrac{q_{1i}}{q_{0i}}$ is the relative quantity change between time 0 and time 1 , and

$x = \dfrac{q_{2j}}{q_{1j}} \div \dfrac{q_{2i}}{q_{1i}}$ is the relative quantity change between time 1 and time 2 .

It is obvious that z and x are either both larger or both smaller than one, which means that relative quantities do not "bounce" in time 1.

References

Allen, R.G.D. [1975], *Index Numbers in Theory and Practice*, Aldine Publishing Company.

Bortkiewicz, L. [1924], ''Zweck und Structur Einer Preisindexzahl'', *Nordic Statistical Journal*.

Diewert, W.E. [1983], ''The Theory of the Cost of Living Index'', *Price Level Measurement: Proceedings From a Conference Sponsored by Statistics Canada*.

Divisia, F. [1925], ''L'indice monétaire et la théorie de la monnaie'', *Revue d'Economie Politique*, Vol. 39.

Forsyth, F. [1978], ''The Practical Construction of a Chain Price Index Number'', *The Journal of the Royal Statistical Society, Series A*, Vol. 141, Part 3.

__________ and R. Fowler [1981], ''The Theory and Practice of Chain Price Index Numbers'', *The Journal of the Royal Statistical Society, Series A*, Vol. 144, Part 2.

Frisch, R. [1936], ''Annual Survey of General Economic Theory: The Problem of Index Numbers'', *Econometrica*, Vol. 4, No. 1, January 1936.

Généreux, P.A. [1983], ''Impact of Alternate Formulae on the Consumer Price Index'', *Price Level Measurement: Proceedings From a Conference Sponsored by Statistics Canada*.

Marshall, A. [1887], ''Remedies for Fluctuations of General Prices'', *The Contemporary Review*, Vol. 51.

Zarnowitz, V. [1961], ''Index Numbers and the Seasonality of Quantities and Prices'', *Hearings before the Subcommittee on Economic Statistics...*, Washington.

ENCHAÎNEMENT DES INDICES DE PRIX

To provide you with a version in the official language of your choice, the French text is preceeded by the English text (p.537) in this publication.

Bohdan J. Szulc
Statistique Canada
Division des prix
Section centrale de recherche

RÉSUMÉ

Les indices en chaîne sont considérés moins biaisés et plus pertinents que les indices directs à panier fixe correspondants. Les arguments avancés habituellement pour justifier cette supériorité ne sont cependant pas très convaincants. Ils portent sur la validité d'indices à panier fixe particuliers qui sont enchaînés, et non pas sur la validité de l'indice en chaîne entier. L'auteur revoit ces arguments et examine des conditions dans lesquelles un indice en chaîne donnera probablement de moins bons résultats que son homologue direct.

La condition essentielle est le « rebondissement » de prix relatifs dans l'intervalle allant du temps de base au temps-cible, ce qui implique une corrélation inter-temporelle négative particulière entre les variations de prix. Cette corrélation est définie en termes généraux à la section 4, précédée, à la section 3, par une décomposition de l'écart entre les indices en chaîne et directs correspondants. Dans ces conditions, et en supposant une corrélation négative entre les variations de prix et de quantités, les indices en chaîne de Laspeyres, Paasche et Sauerbeck (choisis à cause de leur utilisation en pratique) ont tendance à ne pas être moins biaisés que leurs homologues directs, comme on le montre dans les sections 5 à 7. Ceci est particulièrement vrai lorsque les temps d'enchaînement se trouvent proches des points d'inflexion de prix relatifs.

Dans les mêmes circonstances, les temps d'enchaînement se trouvent proches des points d'inflexion de quantités relatives, ce qui, comme on l'explique à la section 8, signifie que les paniers utilisés dans les indices enchaînés s'écartent des combinaisons linéaires de ces

paniers qui sont tirés des temps de base et cible. Alors, cependant, un indice en chaîne n'est pas plus pertinent que son homologue direct.

Les conclusions théoriques et pratiques des constatations ci-dessus sont examinées à la section 8. Dans la réalité, il est peu probable que les indices en chaîne soient beaucoup plus biaisés que les indices directs correspondants, mais il faut faire preuve de prudence lorsque l'on procède à l'enchaînement afin de mesurer des changements de prix qui présentent des variations cycliques corrélées avec les variations de quantités. Cela mis à part, il semble qu'il n'existe pas d'autre alternative pratique valable que l'enchaînement d'indices à panier fixe lorsqu'il faut mesurer des changements de prix sur de longues périodes.

1. Formulation du problème

Le présent exposé examine le bien-fondé de l'enchaînement des indices de prix associés à des séries chronologiques. L'enchaînement est une opération qui produit un indice en chaîne, par opposition à un indice direct. On entend par indice direct un indice $P^D_{t/0}$ qui mesure la variation de prix survenue au cours d'un intervalle de temps (0,t) d'un seul coup du temps de base 0 au temps-cible (temps d'observation) t. Un indice en chaîne $P^C_{t/0}$ exprime cette variation au moyen d'un produit d'indices $P_{1/0}$, $P_{2/1}$,..., $P_{t/t-1}$ qui quantifient les variations de prix survenues au cours de tous les sous-intervalles (0,1), (1,2),..., (t – 1,t) résultant de la subdivision de l'intervalle considéré:

$$P^C_{t/0} = \prod_{j=1}^{t} P_{j/j-1} \qquad (1)$$

Les indices en chaîne servent de substituts aux indices directs, notamment lorsque ces derniers se révèlent peu pratiques ou inadéquats, comme c'est le cas lorsque les indices portent sur de très longues périodes. Quand l'écart devient très grand entre le temps de base et le temps-cible, il n'existe aucune formule d'indice direct qui soit satisfaisante, en particulier aucune formule à panier fixe ou formule équivalente[1]. Dans ces conditions, les indices de Laspeyres et de Paasche sont susceptibles non seulement de s'écarter l'un de l'autre, mais aussi de mal décrire la réalité, sinon de perdre toute signification. Le choix d'un panier intermédiaire n'est d'aucun secours, car dans le cas de très longs intervalles

(0,t), le panier proviendra nécessairement d'une période éloignée du temps 0 ou du temps t, ou des deux.

Lorsque les comparaisons directes se limitent à des sous-intervalles, comme c'est le cas dans un indice en chaîne, il est probable que l'écart entre les indices de Laspeyres et de Paasche correspondants soit moins grand, de sorte qu'on peut faire avec plus de confiance un choix arbitraire entre les deux (ou encore décider d'en faire la moyenne) que dans le cas des indices directs portant sur une longue période. En outre, et c'est là un facteur encore plus important, l'idée d'une mise à jour des paniers est en soi très attrayante, car elle permet de calculer des indices de prix plus conformes à la réalité et offre un moyen de comparaison à long terme qui prend la forme d'une suite très naturelle de mesures qui se corrigent d'elles-mêmes.

A. Marshall [1887] fut le premier auteur à proposer une mise à jour périodique des pondérations et un enchaînement des indices, de façon à permettre des comparaisons valides de prix à travers le temps. Une élégante justification théorique du procédé de l'enchaînement a été fournie par F. Divisia [1925] qui définissait des indices de prix et de quantités pour des sous-intervalles de temps d'une façon différentielle. On obtient alors l'indice en chaîne par l'intégration de ces indices tout au long de la période. Divisia a souligné que la formule de l'indice ne pose virtuellement aucun problème lorsque les sous-intervalles sont infinitésimaux ou, en pratique, très courts. Le principe de l'enchaînement a également reçu la faveur de plusieurs auteurs contemporains s'intéressant à la théorie de l'indice du coût de la vie. Selon eux, les « vrais » indices du coût de la vie, qu'il est malheureusement impossible de réaliser en pratique, peuvent être approchés de manière satisfaisante par des indices de prix à panier fixe de Laspeyres et de Paasche dont le calcul est possible, en particulier lorsque ces derniers sont voisins l'un de l'autre. De l'avis de W. Diewert [1983, section 4], « Il en résulte que nous devrions construire les indices selon le principe de l'enchaînement plutôt qu'en fonction d'une base fixe ».

Compte tenu de l'ensemble ou d'une partie de ces arguments théoriques et pratiques, ceux qui construisent les indices appliquent souvent le procédé de l'enchaînement lors des calculs empiriques. Il faut toutefois éviter de faire appel à ce procédé sans discernement, car dans certaines circonstances les indices en chaîne ne mesurent pas mieux l'évolution des prix que leurs homologues directs, et en donnent même une image plus fausse. Dans

le présent exposé, ces circonstances sont analysées pour des indices de prix à panier fixe[2]; ceux-ci ont été choisis en raison de leur utilisation fréquente en pratique et de leurs propriétés arithmétiques transparentes. Il est aussi possible d'effectuer une analyse similaire pour tous les indices de prix qui prennent la forme de moyennes arithmétiques d'indices de sous-agrégats, quoique l'interprétation des résultats ne soit pas toujours facile.

2. Propriétés arithmétiques des indices en chaîne

On peut facilement déduire, d'après sa formule, les caractéristiques de base d'un indice à panier fixe. Ce qu'il importe avant tout d'observer, c'est qu'un tel indice, calculé pour un agrégat donné de produits, peut toujours s'interpréter comme une moyenne arithmétique pondérée des indices de prix des sous-agrégats correspondants. Tout indice à panier fixe possède donc l'ensemble des caractéristiques d'une moyenne arithmétique[3], ce qui facilite énormément la compréhension de son comportement. Par exemple, aucun indice à panier fixe portant sur un agrégat donné ne peut atteindre une valeur supérieure (inférieure) à la valeur du plus élevé (moins élevé) des indices de prix des sous-agrégats, et on peut s'attendre à ce qu'un indice à panier fixe donné soit particulièrement sensible aux variations de prix des produits dont la pondération est élevée, c.-à-d. ceux qui représentent une part importante du panier (ces parts étant évaluées en fonction des prix au temps de base).

Les caractéristiques d'un indice de prix en chaîne, par contre, demeurent obscures même lorsque tous les indices enchaînés $P_{j/j-1}$ sont construits suivant la formule d'un indice à panier fixe. Dans ce dernier cas, bien entendu, la connaissance de ces formules peut aider à comprendre le comportement des indices enchaînés particuliers $P_{1/0}, P_{2/1}, \ldots, P_{t/t-1}$, mais elle ne fournit aucune indication apparente quant au comportement de leur produit, c.-à-d. de l'indice en chaîne résultant. En général, un tel indice en chaîne calculé pour un agrégat donné de produits ne peut être interprété comme une moyenne arithmétique des indices de prix des sous-agrégats correspondants (à moins qu'on ne permette des pondérations négatives). Les indices en chaîne peuvent, par conséquent, donner des résultats surprenants, comme on peut le voir dans cet exemple numérique très simple.

Exemple 1

Soient p_0, p_1, p_2, p_3, p_4 et q_0, q_1, q_2, q_3, q_4 les prix et les quantités de trois produits A, B, C aux temps 0, 1, 2, 3, 4, respectivement.

Produit	Prix					Quantité				
	p_0	p_1	p_2	p_3	p_4	q_0	q_1	q_2	q_3	q_4
A	3	4	5	4	3	14	12	10	12	14
B	5	5	5	5	5	10	10	10	10	10
C	7	6	5	6	7	6	8	10	8	6

Les indices de prix à panier fixe de Laspeyres[4] pour les sous-intervalles consécutifs (0,1), (1,2), (2,3) et (3,4) sont les suivants:

$$P^L_{1/0} = \frac{142}{134} \approx 1.06 \qquad P^L_{2/1} = \frac{150}{146} \approx 1.03$$

$$P^L_{3/2} = \frac{150}{150} = 1.00 \qquad P^L_{4/3} = \frac{142}{146} \approx 0.97$$

Leur produit, c.-à-d. l'indice en chaîne de Laspeyres pour l'intervalle entier (0,4) est $P^{CL}_{4/0} \approx 1.06$. Il indique donc une augmentation des prix d'environ 6%, bien que tous les prix soient revenus, au temps 4, au même niveau qu'au temps de base 0.

L'exemple 1 peut paraître artificiel, mais les séries réellement publiées contiennent également certains indices de prix en chaîne aux propriétés déroutantes. On peut mentionner par exemple le cas d'un indice dont la valeur est inférieure à celle des indices de ses deux sous-composantes[5].

Il existe un moyen d'éclairer, dans une certaine mesure, le comportement d'un indice en chaîne, en établissant une comparaison avec un indice direct dont on connaît la formule, et donc les caractéristiques de base. On peut notamment procéder de la sorte lorsque tous les indices enchaînés $P_{j/j-1}$ sont établis selon une seule formule, disons la formule *, comme cela se fait souvent en pratique. L'indice en chaîne $P^{C*}_{t/0}$ obtenu par la multiplication de ces indices $P^{*}_{1/0}$, $P^{*}_{2/1}, \ldots, P^{*}_{t/t-1}$ peut alors être comparé à l'indice direct $P^{D*}_{t/0}$, calculé suivant la même formule *. En général, les deux indices ne sont pas égaux:

$$P^{C*}_{t/0} \neq P^{D*}_{t/0} \tag{2}$$

L'énoncé (2) ne vaut pas pour les formules d'indice dans lesquelles le panier demeure fixe pour tous les indices de prix, quels que soient le temps de base et le temps-cible[6]. Ces formules sont en effet transitives, ce qui signifie précisément que les indices en chaîne et direct correspondants sont toujours identiques. Le présent exposé ne traite pas de cette situation, d'une part en raison de son caractère trivial et d'autre part parce qu'un indice en chaîne construit sans mises à jour du panier semble contredire la raison d'être même de l'enchaînement.

Il n'est pas possible, en ne tenant compte que de l'énoncé (2), de porter un jugement en faveur ou en défaveur de la procédure d'enchaînement, même dans les cas où les séries correspondant à l'indice en chaîne et à l'indice direct présentent un écart considérable et systématique (R. Frisch [1936] parlait de « drift », c.-à-d. « dérive »). On peut toutefois réaliser certains progrès en analysant les facteurs et les conditions qui expliquent l'écart observé entre les indices en chaîne et direct correspondants.

Il peut paraître surprenant de constater le nombre restreint d'articles qui ont été écrits sur le sujet, étant donné l'utilisation très répandue du procédé d'enchaînement des indices. Cela étant dit, les méthodes et les idées présentées ici ont été traitées auparavant par plusieurs auteurs. Mentionnons en premier lieu L. Bortkiewicz [1924], qui a élaboré une méthode de décomposition de l'écart observé entre deux indices de prix (plus précisément les indices de Paasche et de Laspeyres) et dont il est fait ici abondamment usage dans un autre contexte. R. Frisch [1936] a formulé le problème de base qui fait l'objet du présent exposé,

mais a fait une analyse détaillée seulement pour la formule de Sauerbeck et n'a pas examiné toutes les conséquences. V. Zarnowitz [1961] a étudié le problème sous l'angle de la saisonnalité des prix et des quantités, mais il a soulevé implicitement certaines questions générales pour la première fois. R. Allen [1975] a été le premier à étendre l'analyse à plus de deux indices enchaînés, mais il ne s'est intéressé qu'à la formule de Laspeyres. Enfin, les articles de F. Forsyth [1978] et de F. Forsyth et R. Fowler [1981] comptent parmi les contributions les plus récentes à l'étude du sujet.

3. Facteurs de l'écart entre les indices correspondants en chaîne et direct à panier fixe

L'écart entre les indices correspondants en chaîne et direct à panier fixe direct en chaîne est examiné dans cette section d'un point de vue très général, quelle que soit la formule particulière utilisée.

Soit $p^{C*}_{t/0}$ un indice de prix en chaîne qui résulte de l'enchaînement d'indices à panier fixe dont le j^e est défini de la façon suivante:

$$P^*_{j/j-1} = \frac{\Sigma\, p_j q_{k_j}}{\Sigma\, p_{j-1} q_{k_j}} = \frac{\Sigma\, (p_j/p_{j-1})\,(p_{j-1} q_{k_j})}{\Sigma\, (p_{j-1} q_{k_j})} \tag{3}$$

où

p_j et p_{j-1} sont les prix d'un produit particulier aux temps j et j – 1 respectivement,
q est la quantité fixe attribuée à ce produit dans ce j^e indice enchaîné, selon le panier du temps de référence k_j, et
Σ est le symbole de sommation sur tous les produits que renferme l'agrégat donné.

Autrement dit:

$$P^{C*}_{t/0} = \prod_{j=1}^{t} \frac{\Sigma\, (p_j/p_{j-1})\,(p_{j-1} q_{k_j})}{\Sigma\, (p_{j-1} q_{k_j})} \tag{4}$$

De même, soit $p^{D*}_{t/0}$ un indice de prix direct à panier fixe que l'on définit de la façon suivante:

$$P^{D*}_{t/0} = \frac{\Sigma\ p_t q_d}{\Sigma\ p_0 q_d} \tag{5}$$

où

p_t et p_0 sont les prix d'un produit particulier au temps-cible et au temps de base, respectivement, q_d est la quantité fixe attribuée à ce produit dans cet indice direct particulier, selon le panier du temps de référence d, et

Σ est le symbole de sommation sur tous les produits que renferme l'agrégat donné.

L'indice ci-dessus peut s'écrire sous une forme différente:

$$P^{D*}_{t/0} = \prod_{j=1}^{t} \frac{\Sigma\ p_j q_d}{\Sigma\ p_{j-1} q_d} = \prod_{j=1}^{t} \frac{\Sigma\ (p_j/p_{j-1})\ (p_{j-1} q_d)}{\Sigma\ (p_{j-1} q_d)} \tag{6}$$

Le quotient de l'indice en chaîne et de l'indice direct, tels qu'ils sont présentés aux formules (4) et (6), peut s'exprimer sous forme du produit suivant:

$$P^{C*}_{t/0} \div P^{D*}_{t/0} = \prod_{j=1}^{t} F^*_j \tag{7}$$

où les facteurs F^*_j correspondent aux quotients de deux moyennes arithmétiques des mêmes rapports de prix, mais pondérées différemment:

$$F^*_j = \frac{\Sigma\ (p_j/p_{j-1})\ (p_{j-1} q_{k_j})}{\Sigma\ (p_{j-1} q_{k_j})} \div \frac{\Sigma\ (p_j/p_{j-1})\ (p_{j-1} q_d)}{\Sigma\ (p_{j-1} q_d)} \tag{8}$$

La proportion entre les deux pondérations attribuées à un produit donné dans les moyennes ci-dessus est comme suit:

$$(p_{j-1} q_{k_j}) \div (p_{j-1} q_d) = q_{k_j}/q_d$$

Elle est donc égale au rapport entre les quantités de ce produit présentes dans les paniers fixes aux temps de référence k_j et d, respectivement. À des fins de simplication, désignons par y ces rapports de quantités:

$$y = q_{k_j}/q_d \tag{9}$$

et par x les rapports de prix dont on fait la moyenne:

$$x = p_j/p_{j-1} \tag{10}$$

Chacun des facteurs F_j^* se prête maintenant à la décomposition suivante, semblable à celle qu'a proposée Bortkiewicz [1924][7]:

$$F_j^* = 1 + r_{xy} \cdot V_x \cdot V_y \quad \text{pour } j = 1, 2,\ldots,t \tag{11}$$

où

r_{xy} est le coefficient de corrélation linéaire entre les rapports de prix x et les rapports de quantités y

V_x est le coefficient de variation des rapports de prix x

V_y est le coefficient de variation des rapports de quantités y.

Il découle de (7) que l'indice en chaîne $P_{t/0}^{C*}$ est supérieur (inférieur) à l'indice direct correspondant $P_{t/0}^{D*}$ si les facteurs F_j^* sont de façon prédominante supérieurs (inférieurs) à l'unité, ce qui, en vertu de (11), dépend uniquement du signe des coefficients de corrélation r_{xy}[8].

4. Corrélation entre les variations de prix courantes et les variations de quantités cumulatives

Pour établir dans quelle direction l'indice en chaîne $P^{C*}_{t/0}$ dévie par rapport à son homologue direct $P^{D*}_{t/0}$, il suffit maintenant de déterminer les signes des coefficients de corrélation r_{xy} dans le cadre de l'ensemble des produits que renferme l'agrégat donné. Chaque coefficient exprime la corrélation qui existe entre les rapports de prix $x = p_j/p_{j-1}$ et les rapports de quantités $y = q_{k_j}/q_d$. Les rapports de prix x mesurent les variations de prix **courantes** qui surviennent au cours de sous-intervalles consécutifs (j – 1,j) précédant immédiatement les temps d'enchaînement j. Les rapports de quantités y, en revanche, mesurent les variations de quantités qui **s'accumulent** au cours des intervalles de temps (d, k_j) ou (k_j, d), c.-à-d. entre le temps de référence du panier associé à l'indice direct $P^{D*}_{t/0}$ (une constante) et les temps de référence des paniers associés aux indices enchaînés $P^{C*}_{j/j-1}$ (en général une variable). Comme les x et les y ne se rapportent pas au même intervalle de temps, il n'est pas facile de faire des hypothèses quant au signe qui prédominera dans les coefficients r_{xy} dans une situation donnée. En fait, certains auteurs se montrent hésitants face à cette question (Allen [1975, pp.197-199]), tandis que d'autres n'en font voir qu'une seule facette (Frisch [1936, p.9], dans des commentaires sur les indices de Laspeyres et de Paasche).

Il est possible de surmonter en partie les difficultés mentionnées en analysant d'autres coefficients de corrélation, plus faciles à interpréter. La méthode consiste à tirer parti d'une relation particulière qui vaut pour toute triade de coefficients de corrélation r_{xy}, r_{xz} et r_{yz} calculés pour une même population. Cette relation permet d'établir des limites pour r_{xy} en fonction de r_{xz} et de r_{yz} de la façon suivante[9]:

$$r_{xz} \cdot r_{yz} - [(1-r^2_{xz})\,(1-r^2_{yz})]^{1/2} \le r_{xy} \le r_{xz} \cdot r_{yz} + [(1-r^2_{xz})\,(1-r^2_{yz})]^{1/2} \qquad (12)$$

La double inégalité (12) n'est pas un moyen très efficace de déduire la valeur de r_{xy}, en particulier lorsque les corrélations entre x et z et entre y et z sont faibles l'une et l'autre. Par contre, si l'une des deux corrélations devient plus prononcée, les limites de r_{xy} tendent

à se rapprocher et, en particulier, le signe de r_{xy} a tendance à être déterminé par la combinaison des signes de r_{xz} et r_{yz}. Lorsque r_{xz} et r_{yz} sont du même signe, r_{xy} a tendance à être positif plutôt que négatif, tandis que l'inverse se produit lorsque r_{xz} et r_{yz} sont de signes contraires. La connaissance des signes de r_{xz} et r_{yz} se révèle donc utile pour formuler des hypothèses quant au signe de r_{xy}.

Définissons la variable z de la façon suivante:

$$z = p_{k_j}/p_d \qquad (13)$$

Le coefficient r_{yz}, la variable y étant définie par (9), mesure alors la corrélation entre les variations de quantités $y = q_{k_j}/q_d$ et les variations de prix $z = p_{k_j}/p_d$ de produits particuliers au cours du même intervalle de temps (d, k_j). Quel que soit l'intervalle, ces coefficients ont une tendance à être négatifs pour les biens de consommation, en raison du phénomène de substitution qu'entraîne une variation de prix. Par conséquent, en vertu de (12), les signes des coefficients r_{xy}, ceux qui nous intéressent, dépendent dans une grande mesure des signes des coefficients r_{xz} qui expriment la corrélation inter-temporelle entre les variations de prix courantes $x = p_j/p_{j-1}$ et les variations de prix $z = p_{k_j}/p_d$ qui s'accumulent entre le temps d (fixe) et le temps k_j (mobile). Si les r_{xz} ont tendance à être négatifs pour $j = 1,2,...,t$, les r_{xy} seront plutôt positifs et l'indice en chaîne $P^{C*}_{t/0}$ aura tendance à être supérieur à l'indice direct correspondant $P^{D*}_{t/0}$. L'inverse est vrai si les r_{xz} ont tendance à être positifs.

La corrélation entre x et z est une corrélation inter-temporelle entre des variations de prix [10]. Le signe du coefficient r_{xz} dépend donc de la façon dont ont évolué les prix au cours des périodes à l'étude. Il dépend également du choix des temps j, k_j et d qui ont servi au calcul de l'indice en chaîne et de l'indice direct qu'on veut comparer. Les temps $j = 1,2,...,t$ sont les limites des sous-intervalles (0,1), (1,2),..., (t - 1,t) selon lesquels l'intervalle de temps (0,t) a été divisé. Ils correspondent aux moments où les indices consécutifs $P^*_{j/j-1}$ sont enchaînés aux suivants. Les temps k_j (multiples, en règle générale) et d (unique) sont les temps de référence des paniers fixes associés aux indices enchaînés $P^*_{j/j-1}$ et à l'indice direct $P^{D*}_{t/0}$, respectivement. Ces temps dépendent de la formule de l'indice, si bien que pour examiner la corrélation inter-temporelle entre les variations

de prix, il faut tenir compte de la formule choisie. C'est ce qu'on fera dans les deux prochaines sections avec les formules de Laspeyres et de Paasche, retenues en raison de leur importance dans la théorie et les applications des indices.

La mesure dans laquelle le signe de r_{xz} influe sur le signe de r_{xy} dépend de r_{yz}, et son impact sera réduit si la corrélation entre y et z est faible. Comme nous l'avons mentionné ci-dessus, les r_{yz} sont plutôt négatifs pour les biens de consommation, mais la corrélation est rarement très forte. En fait, la substitution à cause des prix n'est pas le seul facteur qui fait varier la consommation (voir P. Généreux [1983]). Il existe toutefois des procédés d'enchaînement qui font intervenir des mécanismes imitant un phénomène très poussé de substitution attribuable aux prix. Dans ce cas, le signe inverse de celui de r_{xz} est nécessairement transmis à r_{xy}. C'est le cas de l'enchaînement avec la formule de Sauerbeck, dont on discutera après ceux des formules de Laspeyres et de Paasche.

5. Corrélation inter-temporelle entre les variations de prix: indice en chaîne de Laspeyres

Dans le cas des indices de Laspeyres, les formules comparées sont:

$$P^{D*}_{t/0} = P^{DL}_{t/0} = \frac{\Sigma\, p_t q_0}{\Sigma\, p_0 q_0} \qquad \text{et}$$

$$P^{C*}_{t/0} = P^{CL}_{t/0} = \prod_{j=1}^{t} \frac{\Sigma\, p_j \; q_{j-1}}{\Sigma\, p_{j-1} q_{j-1}}, \qquad \text{de sorte que}$$

$$q_{k_j} = q_{j-1} \text{ et } q_d = q_0 \text{ , et que}$$

$$x = p_j/p_{j-1},\ y = q_{j-1}/q_0 \text{ et } z = p_{j-1}/p_0$$

Il est à remarquer que pour j = 1, on obtient $q_{k_j} = q_d$, d'où, selon (8), il ressort que $F^L_1 = 1$. Pour tous les autres j, comme on l'a vu à la section précédente, les F^L_j ont tendance à dépendre du signe des coefficients r_{xz} correspondants. Dans le cas des indices de Laspeyres le coefficient r_{xz} correspondant à un F^L_j donné mesure la corrélation entre les variation

de prix courantes p_j/p_{j-1} et les variations de prix p_{j-1}/p_0 accumulées depuis le temps de base 0. Autrement dit, cette corrélation se rapporte aux variations de prix survenues au cours des paires d'intervalles (j - 1, j) et (0,j - 1). Pour j = 2,3,...,t, ces paires sont les suivantes:

j	Intervalles correspondant à		
	p_{j-1}/p_0	et	p_j/p_{j-1}
2	(0,1)	et	(1,2)
3	(0,2)	et	(2,3)
		.	
		.	
		.	
t-1	(0,t-2)	et	(t-2,t-1)
t	(0,t-1)	et	(t-1, t)

Examinons le comportement des variations des prix relatifs entre le temps de base 0 et le temps-cible t. Si ces variations ont gardé la même tendance pendant toute la période [11], c.-à-d. si l'ordre des produits en termes de leurs prix relatifs est demeuré le même pendant toute la période, tous les r_{xz} seront positifs pour j = 2,3,...,t. En supposant que les r_{yz} soient négatifs, il s'ensuit que l'indice en chaîne de Laspeyres aura tendance à être inférieur à son homologue direct. Cette tendance devient moins prononcée si les enchaînements sont moins fréquents.

Par contre, si entre le temps de base 0 et le temps-cible t les prix relatifs ont tendance à « rebondir », certains r_{xz} seront négatifs et l'indice en chaîne de Laspeyres aura tendance à excéder l'indice direct correspondant. Si seulement deux indices sont enchaînés, l'effet des rebondissements sera d'autant plus fort que le temps 0 et le temps t seront voisins des crêtes des variations des prix relatifs et que les points d'enchaînement j seront voisins des creux (ou vice versa). L'addition de points d'enchaînement à l'intérieur d'un même cycle d'évolution des prix relatifs aura en général pour effet de réduire l'impact des rebondissements.

Lorsque les variations des prix relatifs présentent des cycles répétitifs et que les temps d'enchaînement sont choisis là où se trouvent les crêtes et les creux consécutifs (ou à proximité), l'indice en chaîne de Laspeyres peut s'écarter indéfiniment de son homologue direct. La synchronisation des phases d'enchaînement avec celles des variations des prix relatifs exerce sur l'indice en chaîne un effet qui s'apparente à celui de la synchronisation des phases du vent avec les mouvements périodiques de la structure du pont suspendu Tacoma Narrows, qui s'est effondré en 1940.

Voici un exemple numérique où la tendance des prix relatifs se maintient et où il y a une corrélation négative entre les variations de prix et les variations de quantités:

Exemple 2

Produit	Prix					Quantité				
	p_0	p_1	p_2	p_3	p_4	q_0	q_1	q_2	q_3	q_4
A	3	4	5	6	7	14	12	10	8	6
B	5	5	5	5	5	10	10	10	10	10
C	7	6	5	4	3	6	8	10	12	14

Les indices de prix à panier fixe de Laspeyres pour les sous-intervalles de temps consécutifs sont les suivants:

$$P^L_{1/0} = \frac{142}{134} \approx 1.06 \qquad P^L_{2/1} = \frac{150}{146} \approx 1.03$$

$$P^L_{3/2} = \frac{150}{150} = 1.00 \qquad P^L_{4/3} = \frac{142}{146} \approx 0.97$$

Leur produit, c.-à-d. l'indice en chaîne de Laspeyres pour l'intervalle entier (0,4), est $P^{CL}_{4/0} \approx 1.06$, ce qui est inférieur à l'indice direct de Laspeyres correspondant, dont la valeur est $P^{DL}_{4/0} = \frac{166}{134} \approx 1.24$. La différence s'amoindrit si un seul enchaînement (au lieu de trois) est effectué, par exemple au temps 2. Dans ce dernier cas, les indices de Laspeyres

enchaînés sont $P^{L}_{2/0} = \frac{150}{134} \approx 1.12$ et $P^{L}_{4/2} = \frac{150}{150} = 1.00$, ce qui donne un indice en chaîne de Laspeyres $P^{CL}_{4/0} \approx 1.12$.

L'exemple 1 (section 2) présente un cas où les prix relatifs « rebondissent » (leur tendance s'inversant au temps 2), et où il y a une corrélation négative entre les variations de prix et de quantités. S'il n'y a qu'un seul enchaînement d'indices, au temps 2, les indices de Laspeyres enchaînés $P^{L}_{2/0} = \frac{150}{134} \approx 1.12$ et $P^{L}_{4/2} = \frac{150}{150} = 1.00$ donnent un indice en chaîne de Laspeyres $P^{CL}_{4/0} \approx 1.12$, supérieur à l'indice direct de Laspeyres correspondant $P^{DL}_{4/0} = 1.00$. La différence s'amenuise, bien qu'elle demeure du même signe, si l'enchaînement est effectué à trois reprises, aux temps 1, 2 et 3. La valeur de l'indice en chaîne de Laspeyres est alors de 1.06, comme on l'a vu à la section 2. Remarquons que le temps d'enchaînement 1 fait partie de l'intervalle (0,2), au cours duquel les prix relatifs ont maintenu leur tendance. C'est le cas également du temps d'enchaînement 3 qui se trouve à l'intérieur de l'intervalle (2,4). Enfin, s'il n'y a qu'un enchaînement au temps 1, les indices de Laspeyres enchaînés $P^{L}_{1/0} = \frac{142}{134} \approx 1.06$ et $P^{L}_{4/1} = \frac{142}{146} \approx 0.97$ donnent un indice en chaîne $P^{CL}_{4/0} \approx 1.03$, inférieur à l'indice obtenu avec un enchaînement au temps 2, où se produit l'inversion.

L'exemple 3 (section 7) présente un cas où se manifestent des cycles répétitifs dans l'évolution des prix relatifs et l'incidence de ces cycles sur les indices en chaîne.

6. Corrélation inter-temporelle entre les variations de prix: indice en chaîne de Paasche

Dans le cas des indices de Paasche, les formules comparées sont:

$$P^{D*}_{t/0} = P^{DP}_{t/0} = \frac{\Sigma\, p_t q_t}{\Sigma\, p_0 q_t} \quad \text{et}$$

$$P^{C*}_{t/0} = P^{CP}_{t/0} = \prod_{j=1}^{t} \frac{\Sigma\, p_j q_j}{\Sigma\, p_{j-1} q_j}, \text{ de sorte que}$$

$$q_{k_j} = q_j \text{ et } q_d = q_t, \text{ et que}$$

$$x = p_j/p_{j-1},\ y = q_j/q_t \text{ et } z = p_j/p_t$$

On observe que pour $j = t$, on a $q_{k_j} = q_d$, de sorte que $F_t^P = 1$ en vertu de (8). Pour les autres j, comme on l'a vu à la section précédente, les F_j^P ont tendance à dépendre du signe des r_{xz} correspondants, ou du signe inverse des $r_{x\ 1/z}$ correspondants. Ce dernier coefficient mesure la corrélation entre les variations de prix courantes p_j/p_{j-1} et les variations de prix p_t/p_j accumulées jusqu'au temps-cible t. En d'autres termes, cette corrélation se rapporte aux variations de prix qui sont survenues au cours des paires d'intervalles (j-1,j) et (j,t). Pour $j = 1,2,\ldots,t\text{-}1$, ces paires sont les suivantes:

j	Intervalles correspondant à		
	$p_j/p_{j\text{-}1}$	et	p_t/p_j
1	(0,1)	et	(1,t)
2	(1,2)	et	(2,t)
.		.	
.		.	
.		.	
t-2	(t-3,t-2)	et	(t-2,t)
t-1	(t-2,t-1)	et	(t-1,t)

Examinons l'évolution des variations des prix relatifs entre le temps de base 0 et le temps-cible t, comme nous l'avons fait pour les indices de Laspeyres. Si ces variations ont gardé la même tendance pendant la période entière, tous les $r_{x\ 1/z}$ seront positifs pour $j = 1,2,\ldots,t\text{-}1$, de sorte que les r_{xz} seront tous négatifs (puisque ni x ni z ne peuvent normalement prendre de valeurs négatives). En supposant que les r_{yz} soient négatifs, il s'ensuit que l'indice en chaîne de Paasche aura tendance à être supérieur à son homologue direct.

Si les prix relatifs ont tendance à « rebondir » entre le temps de base 0 et le temps-cible t, certains $r_{x\ 1/z}$ seront négatifs, les r_{xz} correspondants seront positifs, et l'indice en chaîne de Paasche aura tendance à être inférieur à son homologue direct. L'effet des rebondissements dépend de la position des temps 0, j et t de la même façon que pour les indices de Laspeyres, traités à la section 5.

Dans l'exemple numérique 2 (section 5), les variations des prix relatifs évoluent suivant une tendance qui se maintient et présentent une corrélation négative avec les variations

de quantités. L'indice direct de Paasche est $P^{DP}_{4/0} = \frac{134}{166} \approx 0.81$. L'indice en chaîne de Paasche avec enchaînement aux temps 1, 2 et 3 est $P^{CP}_{4/0} \approx 1.03 \times 1.00 \times 0.97 \times .093 \approx 0.94$, donc supérieur à l'indice direct. Avec un seul enchaînement au temps 2, l'indice en chaîne de Paasche est $P^{CP}_{4/0} \approx 1.00 \times 0.89 \approx 0.89$, et avec un seul enchaînement au temps 1, l'indice est $P^{CP}_{4/0} \approx 1.03 \times 0.85 \approx 0.87$, c.-à-d. encore une fois supérieur à l'indice direct de Paasche, bien que la différence soit moins prononcée.

Dans l'exemple numérique 1 (section 2), les prix relatifs « rebondissent » au temps 2, et sont en corrélation négative avec les variations de quantités. L'indice direct de Paasche est $P^{DP}_{4/0} = \frac{134}{134} = 1.00$. L'indice en chaîne de Paasche avec enchaînement aux temps 1, 2 et 3 est $P^{CP}_{4/0} \approx 1.03 \times 1.00 \times 0.97 \times 0.94 \approx 0.94$, c'est-à-dire inférieur à l'indice direct. Avec un seul enchaînement au temps 2, l'indice en chaîne de Paasche est $P^{CP}_{4/0} \approx 1.00 \times 0.89 \approx 0.89$ et avec un seul enchaînement au temps 1, l'indice est $P^{CP}_{4/0} \approx 1.03 \times 0.94 \approx 0.97$, encore une fois inférieur à l'indice direct. La différence la plus grande apparaît lorsque l'enchaînement se fait au temps 2 (là où survient l'inversion de l'évolution des prix relatifs).

7. Corrélation inter-temporelle entre les variations de prix: indice en chaîne de Sauerbeck

Prenons maintenant comme exemple une moyenne arithmétique simple (non pondérée) de rapports de prix, qu'on appelle indice de Sauerbeck. Cette formule est parfois utilisée en pratique, mais pas aux niveaux supérieurs d'agrégation; on s'en sert plutôt pour combiner certaines données individuelles sur les prix (recueillies par exemple tous les mois) et construire l'indice de prix d'un mois sur l'autre au niveau d'agrégation le plus bas. Lorsqu'il y a modification de l'ensemble des articles choisis pour représenter l'évolution des prix de l'agrégat du plus bas niveau considéré, ces indices peuvent être enchaînés. En outre, de manière à disposer d'un algorithme uniforme pour le calcul de l'indice, le procédé d'enchaînement peut être étendu à tous les indices de Sauerbeck d'un mois sur l'autre du plus bas niveau, que les articles représentatifs aient changé ou non. L'indice en chaîne de Sauerbeck présente toutefois des caractéristiques bien particulières.

Établissons une comparaison entre l'indice en chaîne de Sauerbeck et l'indice direct correspondant:

$$P_{t/0}^{D*} = P_{t/0}^{DS} = \frac{\Sigma\ (pt/p0)}{N} \text{ et}$$

$$P_{t/0}^{C*} = P_{t/0}^{CS} = \prod_{j=1}^{t} P_{j/j-1}^{S} = \prod_{j=1}^{t} \frac{\Sigma\ (pj/pj\text{-}1)}{N},$$

où N est le nombre d'articles dont on calcule la moyenne des rapports de prix dans chaque indice. Il est possible de transformer les indices directs et en chaîne ci-dessus en indices implicites à panier fixe, comme suit:

$$P_{t/0}^{DS} = \frac{\Sigma\ (p_t/p_0)}{N} = \frac{\Sigma\ (p_t/p_0)(p_0\frac{1}{p_0})}{\Sigma\ (p_0\frac{1}{p_0})} \text{ et}$$

$$P_{j/j-1}^{S} = \frac{\Sigma\ (p_j/p_{j-1})}{N} = \frac{\Sigma\ (p_j/p_{j-1})(p_{j-1}\frac{1}{p_{j-1}})}{\Sigma\ (p_{j-1}\frac{1}{p_{j-1}})}$$

Nous pouvons conclure que les indices directs et en chaîne de Sauerbeck sont équivalents à des indices à panier fixe, dont les quantités fixes implicites sont $q_d = \frac{1}{p_0}$ et $q_{k_j} = \frac{1}{p_{j-1}}$. Autrement dit, au cours de chaque intervalle de temps (0,j-1), la variation des quantités fixes implicites est l'inverse de celle des prix. Par conséquent, les indices de Sauerbeck correspondent à des indices de prix à panier fixe calculés dans l'hypothèse d'une substitution extrêmement forte attribuable aux prix, qui se manifesterait pour tous les produits.

Dans ces circonstances, tous les r_{yz} sont égaux à moins un et chaque r_{xy} est déterminé de manière non équivoque par r_{xz} (avec le signe opposé) ou, ce qui est plus facile à suivre, par $r_{x1/z}$ (avec le même signe). Ce dernier coefficient mesure la corrélation entre les variations de prix courantes p_j/p_{j-1} et les variations de prix cumulatives p_{j-1}/p_0. Ces dernières ont trait aux intervalles (j-1,j) et (0,j-1), les mêmes que dans le cas des indices de Laspeyres.

Si la même tendance des variations des prix relatifs se maintenait pendant toute la période qui s'écoule entre le temps de base 0 et le temps-cible t, tous les r_{xy} seraient négatifs. L'indice en chaîne de Sauerbeck serait alors inférieur à son homologue direct. En revanche, c'est le contraire qui est vrai lorsque les prix relatifs ont tendance à « rebondir ». Sous cet angle, la relation entre les indices directs et en chaîne de Sauerbeck est analogue à celle qu'on avait établie pour les indices de Laspeyres. Toutefois, pour une même évolution des variations des prix relatifs, l'impact sera généralement beaucoup plus grand sur l'indice en chaîne de Sauerbeck. S'il y a des rebondissements répétitifs et que le cycle d'enchaînement est en phase avec celui du mouvement des prix relatifs, on pourra obtenir des résultats surprenants. On le constate en examinant l'exemple numérique qui suit, dans lequel les rapports de prix ont été calculés pour deux articles seulement.

Exemple 3

Article	Prix					Rapport de prix			
	p_0	p_1	p_2	p_3	p_4	p_1/p_0	p_2/p_1	p_3/p_2	p_4/p_3
A	1	2	1	2	1	2.0	0.5	2.0	0.5
B	2	1	2	1	2	0.5	2.0	0.5	2.0
Moyenne simple						1.25	1.25	1.25	1.25

Calculons l'indice en chaîne de Sauerbeck:

$$P_{0/4}^{CS} = \prod_{j=1}^{4} \frac{\Sigma\,(p_j/p_{j-1})}{2} = 1.25^4 \approx 2.44.$$

Il indique une progression des prix de 144%, bien que les prix soient revenus au temps 4 au même point qu'au temps 0. Il est à remarquer que si les indices enchaînés avaient été définis explicitement sous forme d'indices à panier fixe avec une quantité « un » associée à chaque article (où, de façon équivalente, sous forme de rapports de moyennes arithmétiques simples de prix), l'indice en chaîne aurait été égal à l'unité:

$$\prod_{j=1}^{4} \frac{\Sigma\, p_j \cdot 1}{\Sigma\, p_{j-1} \cdot 1} = \left(\frac{3}{3}\right)^4 = 1$$

8. Bien-fondé des indices en chaîne

On ne peut examiner le bien-fondé des indices en chaîne sans se demander s'ils sont susceptibles de présenter les caractéristiques qu'on recherche principalement lorsqu'on applique le procédé d'enchaînement, et dans quelles conditions. L'écart entre les indices en chaîne et les indices directs, s'il en existe un, ne constitue pas en général un critère suffisant pour formuler un jugement. Lorqu'on a recours en pratique à l'enchaînement, c'est habituellement parce qu'on juge que les indices directs sont au mieux des mesures incertaines de la variation des prix; si ce n'était pas le cas, l'enchaînement serait inutile. En conséquence, le fait qu'un indice en chaîne donné s'écarte de l'indice direct correspondant n'est pas, en soit, un symptôme de son « inexactitude » ou « exactitude ». Dans certains cas, cependant, il est non seulement possible mais souhaitable de juger de l'inexactitude d'un indice en chaîne donné en se fondant sur la direction dans laquelle il s'écarte de l'indice direct. Un tel jugement, pour être bien étayé, devrait tenir compte de l'évaluation du biais de l'indice direct correspondant.

Le terme « biais » désigne l'écart entre le résultat produit par un indice donné et celui que devrait produire l'indice théoriquement exact. Dans le cas des prix à la consommation, les indices du coût de la vie constituent de tels repères. En raison de l'effet de substitution, on peut faire les énoncés suivants:

- les indices du coût de la vie de type Laspeyres ont tendance à être inférieurs aux indices de prix à panier fixe de Laspeyres correspondants et, en ce sens, on considère que ces derniers présentent un biais par excès;
- les indices du coût de la vie de type Paasche ont tendance à être supérieurs aux indices de prix à panier fixe de Paasche correspondants et, en ce sens, on considère que ces derniers présentent un biais par défaut.

Plus l'intervalle de temps est long, plus les changements de la structure des prix et des modes de consommation susceptibles de se produire sont appréciables, ouvrant ainsi la voie aux

effets de substitution. Par conséquent, les biais par excès et par défaut respectifs des indices de prix à panier fixe de Laspeyres et de Paasche ont également tendance à s'accroître à mesure que le temps-cible s'éloigne du temps de base.

Lorsqu'on a recours à l'enchaînement, des indices à panier fixe sont calculés pour des sous-intervalles de temps relativement courts; il est donc probable que leurs biais demeureront modestes, quelle que soit la distance entre le temps de base et le temps-cible de l'indice en chaîne. En se fondant sur cette prémisse, on prétend très souvent que les indices en chaîne, dans le cas des longs intervalles de temps, sont moins portés à présenter un biais croissant que leurs homologues directs. C'est là une des principales raisons qui justifie l'emploi des indices en chaîne dans la pratique.

Il existe toutefois une faille notable dans un tel raisonnement. Le biais global de l'indice en chaîne entier peut être considérable même si les biais des indices enchaînés particuliers sont peu élevés. Comme on l'a vu aux sections 5 et 6 du présent exposé, une chaîne d'indices à panier fixe de Laspeyres peut donner un résultat supérieur à l'indice direct de Laspeyres correspondant, tandis qu'un indice en chaîne à panier fixe de Paasche peut être inférieur à l'indice direct de Paasche correspondant. En d'autres termes, les indices en chaîne peuvent présenter des biais supérieurs à ceux de leurs homologues directs, ce qui contredit l'opinion générale et la raison d'être de l'enchaînement. Il faut donc éviter d'avoir recours à ce procédé lorsque de tels résultats non souhaitables tendent à se manifester.

Comme on l'a démontré dans les sections précédentes, cela est susceptible de se produire:

- lorsqu'il y a une corrélation négative entre les variations de prix et de quantités, et
- lorsque les prix relatifs « rebondissent » entre le temps de base et le temps-cible.

En pratique, la deuxième condition doit retenir davantage notre attention, car dans la plupart des situations concrètes les variations de prix et de quantités présentent de toute façon une corrélation négative, en raison des substitutions de produits que font les consommateurs à cause des prix. Pour la même raison, la deuxième condition est généralement équivalente à la suivante:

- les quantités relatives « rebondissent » entre le temps de base et le temps-cible[12].

Le biais des indices en chaîne s'accroît lorsque la corrélation devient plus forte entre les variations de prix et de quantités. Il faut donc examiner avec soin les situations où une telle corrélation est susceptible d'exister, en particulier lorsqu'on sait qu'il se produit une très forte substitution à cause des prix. Les variations saisonnières des prix et des quantités en sont un exemple type, tandis qu'il peut n'y avoir aucune corrélation si les variations de prix ou de quantités surviennent au hasard. La formule de Sauerbeck, qui entraîne automatiquement cette corrélation, bien que de façon voilée, devrait incontestablement être évitée dans le procédé d'enchaînement[13].

Le biais des indices en chaîne sera d'autant plus fort que les points d'enchaînement seront voisins des crêtes du mouvement de prix relatifs (et de quantités relatives), et que le temps de base et le temps-cible seront proches des creux de ce même mouvement, ou vice versa. Il est donc essentiel d'éviter une telle synchronisation entre la périodicité d'enchaînement et celle de l'évolution de prix relatifs (ou de quantités relatives). Si l'on ne tient pas compte de cette mise en garde et que la synchronisation se maintient sur plusieurs cycles successifs, le biais de l'indice en chaîne peut s'accroître indéfiniment. Par contre, si l'on ajoute des points d'enchaînement à l'intérieur d'un cycle particulier du mouvement de prix relatifs (ou de quantités relatives), le biais s'amenuisera[14].

En pratique, il est peu probable que les indices en chaîne présenteront des biais beaucoup plus grands que les indices directs correspondants. Cela est particulièrement vrai dans le cas des agrégats composés de nombreux produits, pour les raisons suivantes:

- les variations de prix relatifs (ou de quantités relatives) ont généralement des comportements divers, certaines conservant la même tendance, d'autres ayant tendance à « rebondir »;
- lorsque prix relatifs (quantités relatives) « rebondissent », il est rare que tous les produits suivent le même cycle, de sorte qu'il est presque impossible qu'on ait une synchronisation complète avec le cycle d'enchaînement;
- l'enchaînement est souvent utilisé avec différentes formules et différentes périodicités, à différents niveaux d'agrégation des produits; le résultat final est donc un indice qui est peut-être moins transparent, mais aussi moins vulnérable aux comportements extrêmes que l'indice résultant d'un enchaînement uniforme;

– les indices enchaînés particuliers sont pour la plupart basés sur des paniers fixes correspondant à des moments qui précèdent les points d'enchaînement; les coefficients de corrélation inter-temporelle mesurent donc le lien entre les variations de prix (ou de quantités) survenues au cours d'intervalles de temps disjoints, lien qui se révélera faible selon toute probabilité:

– la substitution attribuable aux prix n'est pas un phénomène qui concerne tous les produits en tout temps; la corrélation négative entre les variations de prix et les variations de quantités est donc normalement faible.

* * *

Une autre des raisons principales qui incite à employer les indices en chaîne dans la pratique, c'est qu'on leur attribue un avantage théorique sur les indices directs, en particulier dans le cas des séries d'indices portant sur de longues périodes. Dans ce cas, des modifications appréciables de la distribution de la consommation, de la production, du commerce, etc. (en termes quantitatifs) ont tendance à se produire entre le temps de base et le temps-cible. Comme aucun panier ne peut correspondre à la fois à ces deux temps, ne serait-ce que de façon approximative, on considère que les indices à panier fixe sont inapplicables, sinon complètement vides de sens, quelle que soit la formule retenue. Dans de telles conditions, le procédé d'enchaînement semble offrir une solution très satisfaisante au problème, en substituant à un indice unique, détaché de la réalité, une chaîne de mesures réalistes des variations de prix.

Cette opinion, largement répandue, se fonde néanmoins sur un raisonnement qui n'est pas sans faille. Le procédé d'enchaînement est conçu de manière à fournir une mesure des variations de prix au cours de l'intervalle entier, depuis le temps de base jusqu'au temps-cible. Les indices de prix relatifs aux sous-intervalles particuliers de cet intervalle ne constituent pas l'objet principal du processus d'enchaînement, bien qu'ils en soient des étapes nécessaires et des sous-produits. En conséquence, le bien-fondé théorique du procédé d'enchaînement doit être examiné à la lumière du comportement de l'indice en chaîne global et non de celui des indices enchaînés particuliers qui servent à son calcul.

Sous cette angle, les indices en chaîne peuvent être considérés comme supérieurs à leurs homologues directs lorsqu'ils assurent un passage en douceur entre le temps de base et le temps-cible, et non pas lorsqu'ils empruntent un détour. Il est logique et avantageux

d'enchaîner les indices de prix aux temps, s'il en est, caractérisés par des distributions quantitatives de la consommation (de la production, du commerce, etc.) qui s'inscrivent dans un mouvement de variations progressives entre le temps de base et le temps-cible. En termes plus rigoureux, ces distributions devraient se rapprocher le plus possible de combinaisons linéaires des distributions existant au temps de base et au temps-cible. À l'inverse, il est illogique d'enchaîner des indices de prix à des temps caractérisés par des distributions quantitatives dont on sait qu'elles s'écartent de façon importante de ce mouvement de variations progressives[15]. Procéder de la sorte équivaudrait à comparer des situations distinctes par l'intermédiaire de situations encore plus différentes, ou, dans le cas extrême, à comparer des situations identiques par l'intermédiaire de situations bien différentes[16].

Les modifications que subit la distribution de la consommation (de la production, du commerce, etc.) selon les produits, qu'elles résultent de changements structurels d'ordre économique ou démographique, d'une évolution de la technologie et des habitudes, ou encore d'autres mutations à long terme, se produisent normalement de façon progressive. Pour tenir compte de telles modifications, il aurait été souhaitable de mettre à jour les paniers fixes et d'enchaîner les indices de prix le plus souvent possible. Toutefois, les distributions quantitatives sont soumises à des variations périodiques, saisonnières et aléatoires, si bien qu'il faudrait exercer avec prudence le choix des paniers fixes associés aux indices enchaînés, ainsi que des points d'enchaînement eux-mêmes. En fait, quoiqu'en dise Divisia, il faut habituellement prévoir une certaine longueur minimale pour les sous-intervalles de temps à partir desquels s'effectuera l'enchaînement.

Cette longueur peut d'ailleurs varier selon le secteur d'activité visé par l'indice. Par exemple, dans le cas des indices de prix relatifs au bétail, on peut tenter de ne pas effectuer l'enchaînement à l'intérieur des cycles des prix des porcins et des bovins. Dans le cas des prix à la consommation, il semble que l'année soit la durée optimale théorique, mais il peut être par ailleurs souhaitable, pour des raisons pratiques, d'espacer davantage les enchaînements[17]. Si l'on augmentait la cadence des enchaînements, on incorporerait dans les indices en chaîne des variations saisonnières et aléatoires des quantités relatives (ainsi que des variations connexes des prix relatifs), bien que ces indices visent à mesurer l'évolution des prix à long terme. Lorsqu'on s'attend à des variations aléatoires appréciables, il pourrait se révéler avantageux de baser les paniers fixes associés aux indices enchaînés

sur des moyennes portant sur plusieurs années consécutives. Si des mises à jour plus fréquentes apparaissent justifiées, on peut utiliser la méthode des moyennes mobiles et ainsi éliminer les perturbations attribuables aux variations aléatoires.

* * *

Il est intéressant, et aussi important, de noter que le sentier à variation progressive des distributions quantitatives est aussi un lieu où il ne se produit pas de « rebondissements » des quantités relatives[18]. En outre, si les variations de prix et de quantités présentent une corrélation, les prix relatifs n'y ont pas, non plus, tendance à « rebondir ». Autrement dit, l'analyse du biais des indices en chaîne et l'examen de leurs avantages théoriques nous amènent à la même mise en garde: la prudence est de rigueur lorsque le procédé d'enchaînement est employé pour mesurer des variations de prix qui présentent des variations cycliques en corrélation avec les variations de quantités.

Malgré ces réserves, il n'y a probablement pas d'autre choix pratique que l'enchaînement lorsqu'il faut établir des séries d'indices sur de longues périodes. Tous les inconvénients des indices en chaîne, notamment leurs propriétés arithmétiques assez singulières et la nécessité de les appliquer avec prudence dans certains cas particuliers, apparaissent minimes en comparaison de ceux des indices directs visant de longs intervalles de temps[19].

Renvois

1 Les indices de prix directs à niveau de vie constant (c.-à-d. les indices directs du coût de la vie) présentent des faiblesses semblables, bien qu'elles ne soient pas aussi évidentes. En fait, les niveaux de satisfaction et les goûts observés varient avec le temps et il y a toujours un indice du coût de la vie de type Laspeyres et de type Paasche.

2 Mutatis mutandis, les mêmes conclusions valent pour les indices de volume (de quantité) à prix fixes, ainsi que pour les comparaisons inter-spatiales.

3 Il est à remarquer que dans le calcul réel des indices qu'on appelle indices à panier fixe, une formule qui équivaut à celle du panier fixe s'applique en général seulement à partir d'un certain niveau d'agrégation des produits (voir, par exemple, le Document de référence de l'indice des prix à la consommation, Statistique Canada, nº 62-553 au catalogue, p.27).

4 Dans le présent document, l'expression « indice de Laspeyres » n'est utilisée qu'au sens strict, c.-à-d. seulement lorsque les quantités fixes sont celles du temps de base de l'indice.

5 On observe dans les séries de l'Indice des prix à la consommation du Canada les indices en chaîne suivants pour mars 1978:

Biens	171.1
Services	171.4
Biens et services	170.8

c.-à-d. que l'indice de l'agrégat est inférieur aux indices des deux composantes (voir Prix à la consommation et indices des prix, Statistique Canada, nº 62-010 au catalogue, octobre-décembre 1978, p.40).

6 Les formules en question sont souvent appelées aujourd'hui « formules de Laspeyres », contrairement au sens strict original (voir le renvoi 4). Ce manque de rigueur dans les appellations est plutôt regrettable. En fait, une telle désignation invalide les énoncés bien connus et fort utiles suivants:

- le produit d'un indice de prix de Laspeyres (Paasche) et de l'indice de volume correspondant de Paasche (Laspeyres) est égal au rapport des valeurs courantes (dépenses);

- un indice de prix (volume) de Fisher est une moyenne géométrique des indices de prix (volume) de Laspeyres et de Paasche correspondants.

En outre, une série d'« indices de Laspeyres » définie avec une telle absence de rigueur:

$$\Sigma\, p_t q_c / \Sigma\, p_0 q_c$$

contient nécessairement un indice de Paasche pour $t = c$, ce qui indique encore une fois combien cette terminologie est regrettable.

7 La décomposition est présentée à l'annexe 1.

[8] Il ne vaut pas la peine d'accorder une attention particulière au cas où l'un ou l'autre de V_x ou V_y est nul. Cette situation correspondrait ou bien à l'absence de variations des prix relatifs au cours de sous-intervalles consécutifs (j-1,j), ou encore à une égalité entre les proportions des produits présents dans tous les paniers ayant servi au calcul des indices enchaînés et les proportions existant dans le panier ayant servi à calculer l'indice direct.

[9] Cette double inégalité peut être déduite directement de la formule du coefficient de corrélation partielle $r_{xy \cdot z}$, en sachant que $r^2_{xy \cdot z} \geq 1$.

[10] On peut définir la variable z d'une autre façon. On peut en effet écrire, au lieu de (13):

$$z = q_j / q_{j-1}$$

Dans ce cas, la corrélation entre y et z est une corrélation inter-temporelle entre des variations de quantités, tandis que la corrélation entre x et z est une corrélation (négative) entre des variations de prix et des variations de quantités pour le même sous-intervalle de temps. Les deux définitions mènent à des conclusions analogues.

[11] S'il n'y a pas de variation des prix relatifs, il n'y a pas non plus de problème d'indice.

[12] Si des quantités relatives présentent une corrélation avec des prix relatifs qui « rebondissent », ils auront aussi tendance à « rebondir ». La condition modifiée peut également être déduite du renvoi 10.

[13] Comme nous l'avons dit précédemment, on utilise parfois la formule de Sauerbeck pour combiner différentes données individuelles sur les prix, recueillies pour des produits particuliers dans des points de vente particuliers. Dans ces cas, il est très probable que les prix relatifs présenteront des « rebondissements ». Si certains prix augmentent moins que d'autres dans un mois ou un trimestre donné, par exemple en raison de rabais spéciaux offerts aux clients, ils auront tendance à monter plus que les autres au cours du mois ou du trimestre suivant.

[14] Ces conclusions découlent directement des formules analysées aux sections 5 et 6. Une simulation du procédé d'enchaînement a été effectuée à la Section centrale de recherche, Division des prix, Statistique Canada. On a utilisé à cette fin des données sur des prix et des quantités présentant des comportements saisonniers divers, mais constants d'une année à l'autre. On a calculé des indices avec un, deux, trois, quatre, six et douze points d'enchaînement par année, en faisant varier les mois d'enchaînement, le cas échéant, ainsi que les mois de base (le temps-cible était toujours fixé à douze mois après le mois de base). Les imprimés et les graphiques produits par l'ordinateur ont confirmé toutes les conclusions théoriques.

[15] Lorsqu'on a besoin d'une mesure de la variation des prix tant à long terme qu'à court terme, il peut y avoir incompatibilité entre les exigences relatives aux deux séries d'indices de prix. Par exemple, si on veut faire une comparaison valable entre deux mois consécutifs, il faut que le panier considéré se rapproche le plus possible des distributions quantitatives de ces mois. Par contre, toute comparaison sur une année entière perdra de sa valeur si l'enchaînement est effectué à certains mois particuliers et que les indices enchaînés sont basés sur des paniers saisonniers. Il faut alors décider laquelle des comparaisons de prix revêt le plus d'importance et fixer les exigences en conséquence. Si elles sont tout aussi importantes l'une que l'autre, on pourra produire une série d'indices d'un mois sur l'autre indépendante de la série des indices d'une année sur l'autre. Malheureusement, la plupart des formules d'indice ne sont pas transitives, de sorte qu'un indice annuel donnera, en général, un résultat différent du produit d'indices mensuels.

[16] Les comparaisons indirectes de prix faites à partir d'une série d'indices de prix de Paasche en sont un bon exemple. Il est possible de prouver (voir annexe 2) qu'elles sont l'équivalent d'un indice en chaîne dans lequel deux années (mois, etc.) consécutives sont comparées par l'intermédiaire de l'année (mois,...) de base de la série initiale d'indices de prix de Paasche. Cela explique peut-être pourquoi de telles comparaisons indirectes sont si fortement réprouvées, se voyant même souvent refuser le nom d'indices de prix. Sinon, il serait difficile de comprendre pourquoi le résultat de la division de deux indices de prix se rapportant à des paniers différents est moins un indice de prix que le résultat de la multiplication de tels indices (c.-à-d. un indice en chaîne ordinaire), aucun d'eux n'étant un indice de prix à panier fixe.

[17] L'enchaînement des indices de prix à la consommation tous les ans exigerait des ressources considérables et pourrait donner des résultats à peine meilleurs que ceux d'un indice en chaîne dont les points d'enchaînement se situeraient, par exemple, tous les trois ans (voir P. Généreux [1983]).

[18] Voir annexe 3.

[19] Dans certains cas, toutefois, les propriétés arithmétiques peuvent être le facteur décisif. Par exemple, il est très avantageux d'avoir des valeurs en termes réels qui sont additives, aussi bien en fonction des produits que du temps. Si c'est ce qu'on recherche avant tout, les indices de prix en chaîne ne sont pas applicables à la déflation de séries en valeurs courantes.

Annexe 1

Bortkiewicz [1924] a montré qu'un écart relatif entre les indices de prix de Paasche et de Laspeyres correspondants peut être transformé de la façon suivante:

$$\left[\frac{\Sigma\ p_t q_t}{\Sigma\ p_0 q_t} - \frac{\Sigma\ p_t q_0}{\Sigma\ p_0 q_0}\right] \div \frac{\Sigma\ p_t q_0}{\Sigma\ p_0 q_0}$$

$$= \left[\frac{\Sigma(p_t/p_0)(p_0 q_t)}{\Sigma(p_0 q_t)} - \frac{\Sigma(p_t/p_0(p_0 q_0)}{\Sigma(p_0 q_0)}\right] \div \frac{\Sigma(p_t/p_0)(p_0 q_0)}{\Sigma(p_0 q_0)}$$

$= r_{xy}\ V_x\ V_y$, où:

$x = p_t/p_0$ est le rapport de prix pour un produit particulier,

$y = (p_0 q_t) \div (p_0 q_0) = q_t/q_0$ est le rapport de quantités pour ce produit,

Σ est le symbole de sommation sur tous les produits (les mêmes pour les deux indices),

V_x est le coefficient de variation des rapports de prix,

V_y est le coefficient de variation des rapports de quantités, et

r_{xy} est le coefficient de corrélation linéaire entre ces rapports de prix et de quantités.

La transformation originale de Bortkiewicz constitue un cas particulier d'une décomposition plus générale d'un écart relatif entre deux moyennes arithmétiques quelconques de la même variable, disons x, mais pondérées selon deux ensembles différents de pondérations, disons m et n. Appelons y le rapport m/n entre ces pondérations.

$$\left[\frac{\Sigma xm}{\Sigma m} - \frac{\Sigma xn}{\Sigma n}\right] \div \frac{\Sigma xn}{\Sigma n} = \left[\frac{\Sigma\ x\frac{m}{n}n}{\Sigma\frac{m}{n}n} - \frac{\Sigma xn}{\Sigma n}\right] \div \frac{\Sigma xn}{\Sigma n} =$$

$$\left[\frac{\Sigma xyn}{\Sigma n} - \frac{\Sigma xn}{\Sigma n} \cdot \frac{\Sigma yn}{\Sigma n}\right] \div \left[\frac{\Sigma xn}{\Sigma} \cdot \frac{\Sigma yn}{\Sigma n}\right] = r_{xy} \cdot \frac{\sigma_x}{\bar{x}} \cdot \frac{\sigma_y}{\bar{y}} = r_{xy}\ V_x\ V_y\ .$$

On suppose que toutes les statistiques sont pondérées avec les pondérations n, qui correspondent à (p_0q_0) dans la transformation originale de Bortkiewicz.

Cette décomposition vaut pour toutes les paires d'indices composites de prix pouvant être présentés sous forme de moyennes arithmétiques du même ensemble de rapports de prix, mais pondérées différemment, comme dans le cas des expressions F^*_j, définies par la formule (8) du texte. La décomposition permet une interprétation intéressante du lien qui existe entre des indices composites de prix et, en particulier, des indices de prix à panier fixe. Il faut noter, toutefois, que cette interprétation n'est valable que si les rapports de prix et de quantités figurant dans les coefficients r_{xy}, V_x et V_y correspondent au niveau le plus bas d'agrégation des produits auquel différentes pondérations sont appliquées lors du calcul des indices composites de prix faisant l'objet de la comparaison. Par définition, les mêmes rapports de prix sont appliqués à ce niveau d'agrégation des produits lors du calcul des deux indices composites de prix.

Si l'on fait les changements nécessaires, la décomposition en question s'applique également à toutes les paires d'indices composites de volume qui peuvent être présentés sous forme de moyennes arithmétiques du même ensemble de rapports de quantités, mais pondérées différemment (donc, à toutes les paires d'indices composites de volume fondés sur des ensembles différents de prix fixes).

Annexe 2

$$\frac{\Sigma\, p_t q_t}{\Sigma\, p_0 q_t} \div \frac{\Sigma\, p_{t-1}\, q_{t-1}}{\Sigma\, p_0 q_{t-1}} = \frac{\Sigma\, p_t q_t}{\Sigma\, p_0 q_t} \cdot \frac{\Sigma\, p_0 q_{t-1}}{\Sigma\, p_{t-1} q_{t-1}}$$

Annexe 3

Soient q_{0i}, q_{1i}, q_{2i} les quantités du produit i consommées (produites, vendues, etc.) aux temps 0, 1, 2 respectivement.
Soient q_{0j}, q_{1j}, q_{2j} les quantités correspondantes du produit j.
Si le temps 1 correspond à une distribution quantitative s'inscrivant dans un mouvement de variations progressives entre le temps 0 et le temps 2, on a:

$$\frac{q_{1j}}{q_{1i}} = \alpha \frac{q_{0j}}{q_{0i}} + (1-\alpha).\frac{q_{2j}}{q_{2i}} \quad , \qquad \text{avec } 0 < \alpha < 1$$

ce qu'on peut aussi écrire de la façon suivante:

$$1 = \alpha \div \left[\frac{q_{1j}}{q_{0j}} \div \frac{q_{1i}}{q_{01}}\right] + (1-\alpha) \cdot \left[\frac{q_{2j}}{q_{1j}} \div \frac{q_{2i}}{q_{1i}}\right]$$

ou, sous une forme plus simple:

$$1 = \alpha \div z + (1-\alpha) \cdot x$$

où:

$z = \dfrac{q_{1j}}{q_{0j}} \div \dfrac{q_{1i}}{q_{0j}}$ est la variation des quantités relatives entre le temps 0 et le temps 1, et

$x = \dfrac{q_{2j}}{q_{1j}} \div \dfrac{q_{2i}}{q_{1i}}$ est la variation des quantités relatives entre le temps 1 et le temps 2.

Il est évident que les variables z et x sont toutes deux soit supérieures, soit inférieures à l'unité, de sorte que les quantités relatives ne « rebondissent » pas au temps 1.

Références

Allen, R.G.D. [1975], *Index Numbers in Theory and Practice*, Aldine Publishing Company.

Bortkiewicz, L. [1924], ''Zweck und Structur Einer Preisindexzahl'', *Nordic Statistical Journal.*

Diewert, W.E. [1983], ''The Theory of the Cost of Living Index'', *La Mesure du Niveau des Prix: Actes du Colloque tenu sous l'égide de Statistique Canada.*

Divisia, F. [1925], ''L'indice monétaire et la théorie de la monnaie'', *Revue d'économie politique*, vol. 39.

Forsyth, F. [1978], ''The Practical Construction of a Chain Price Index Number'', *The Journal of the Royal Statistical Society, Series A*, vol. 141, partie 3.

__________ et R. Fowler [1981], ''The Theory and Practice of Chain Price Index Numbers'', *The Journal of the Royal Statistical Society, Series A*, vol. 144, partie 2.

Frisch, R. [1936], ''Annual Survey of General Economic Theory: The Problem of Index Numbers'', *Econometrica*, vol. 4, n° 1, janvier 1936.

Généreux, P. [1983], ''Impact du choix des Formules sur l'Indice des Prix à la consommation du Canada'' *La Mesure du Niveau des Prix: Actes du Colloque tenu sous l'égide de Statistique Canada.*

Marshall, A. [1887], ''Remedies for Fluctuations of General Prices'', *The Contemporary Review*, vol. 51.

Zarnowitz, V. [1961], ''Index Numbers and the Seasonality of Quantities and Prices'', *Hearings before the Subcommittee on Economic Statistics...*, Washington.

THE TREATMENT OF HOUSING IN A COST-OF-LIVING INDEX: RENTAL EQUIVALENCE AND USER COST

Masako N. Darrough
Faculty of Commerce and Business Administration
University of British Columbia

SUMMARY

This paper investigates the issue of how to account for "housing" in the Consumer Price Index estimation. Housing presents a complex problem due to its long lifetime and its large expenditure share in consumer budgets. Thus, estimates of housing cost have a significant impact on the overall CPI.

Two approaches to the measurement of housing cost are discussed: rental equivalence and user cost. These are conceptually equivalent, but may be quite different operationally. The rental equivalence approach is being experimented with by the United States Bureau of Labor Statistics starting in February 1983. This appears to be a major improvement over the old measure which suffers from serious flaws. In addition, the collection of necessary data has been simplified. Unfortunately, observed market rents and user costs estimated by researchers seem to differ substantially. This large discrepancy is not only puzzling but also problematic in selecting the appropriate method. In order to reconcile the discrepancy, we consider the impact of the Canadian and American Taxes. The major features of the taxes are incorporated to derive user costs and rent.

Two important points are identified. First, the tax structure affects the user cost according to the marginal tax rate of the family. Second, the market rent and the user cost differ due to the differential tax treatment of landlord and owner-occupants. It appears that rent and user costs are inherently different. In addition, if we allow for the possibility of segregated housing markets (between rental and ownership), then the attempt to approximate user cost by market rents is further doomed to failure. Although user cost presents

rather complex estimation problems in terms of defining appropriate variables as well as their collection, the rental equivalence approach lacks sufficient theoretical justification to be used as cost of homeownership, and hence housing.

RÉSUMÉ

Ce texte étudie le problème du traitement de la composante "logement" dans l'indice des prix à la consommation. Lors de l'estimation de l'IPC (qui est utilisé pour un grand nombre de décisions politiques), les biens de consommation durables présentent généralement des problèmes particuliers. Par définition, la consommation des biens durables s'étale sur plus d'une période, par conséquent consommation et achat ne coïncident pas. Le logement, en particulier, constitue un cas plus complexe: d'abord, sa durée est très longue et deuxièmement, sa proportion des dépenses est très importante. Par conséquent, les estimations des coûts du logement ont un effet important sur l'IPC global.

Nous nous centrons sur deux techniques de mesure du coût du logement: l'équivalence locative et le coût d'utilisation. La première pose la question de savoir combien le propriétaire demanderait à quelqu'un (comme lui-même) pour louer sa maison, alors que la seconde se centre sur le coût de la propriété et de l'utilisation de la maison. Ces deux questions sont conceptuellement équivalentes (dans certains environnements), mais elles peuvent être très différentes au point de vue pratique. La technique de l'équivalence locative peut être appliquée par le relevé des loyers réels sur le marché. En fait, le Bureau of Labour Statistics des États-Unis en fait l'expérience à compter de février 1983. Cette technique semble marquer un progrès important par rapport à l'ancienne mesure du logement (spécialement de la possession d'une maison) qui possède des défauts importants (par exemple, la double comptabilisation du prix d'achat et des paiements d'hypothèque, l'omission de la déductibilité des paiements d'intérêt, etc.). Cette technique est très intéressante au point de vue opérationnel, car il est plus simple de collecter les données du marché des loyers que de calculer le coût d'utilisation.

Malheureusement, les loyers recueillis sur le marché et les coûts d'utilisation estimés par les chercheurs semblent être très différents. Cet écart important est non seulement difficile à expliquer mais pose également le problème du choix de la méthode appropriée pour le

calcul de la composante logement de l'IPC. Pour supprimer cette divergence, nous tenons compte de l'effet des taxes sur le coût d'utilisation et sur le marché des loyers. Les principales caractéristiques des systèmes de taxation canadien et américain sont incorporées aux calculs pour fin de comparaison entre les deux mesures. On affirme que les taxes sont importantes et peuvent rendre compte des différences observées entre les loyers et les coûts d'utilisation.

Supposons des marchés parfaits du capital, des loyers et de la revente. En l'absence de taxes et de coûts de transactions, le coût d'utilisation serait fonction du prix d'achat, de la hausse (attendue ou réelle) des prix du logement ainsi que des coûts de la dépréciation et de l'entretien. En situation d'équilibre concurrentiel, on s'attend à ce que les propriétaires transfèrent ces montants à titre de loyers. Lorsque nous incorporons la structure de taxation existante, cependant, cette équivalence ne tient plus. Tous les paiements d'intérêt et de taxe foncière sont déductibles du revenu imposable aux États-Unis. Bien qu'aucun de ces deux paiements ne soit déductible au Canada, l'existence d'un impôt sur le revenu d'investissement a un effet sur le coût d'utilisation. Étant donné que les versements initiaux et les remboursements de prêt hypothécaire ne sont pas imposables, les propriétaires qui se financent personnellement auront en fait des coûts d'opportunité plus bas sur les fonds qu'ils ont investis dans leur maison. De plus, le loyer calculé pour le propriétaire de maison n'est pas imposé alors que les loyers reçus par des propriétaires de maisons locatives le sont.

Deux points importants sont identifiés. Premièrement, la structure d'imposition en général et le taux d'imposition particulier affectent le coût de l'utilisation de la maison pour le ménage concerné. Ainsi, le coût d'utilisation d'une maison dépend de la catégorie d'imposition de la famille. Ceci signifie que le coût d'utilisation est spécifique au ménage. Deuxièmement, la valeur locative sur le marché et le coût d'utilisation d'une maison précise ne sont pas nécessairement les mêmes. Ils peuvent être différents à cause de la différence de traitement par rapport à l'impôt que l'on retrouve entre le propriétaire loueur et le propriétaire occupant.

En résumé, il semble que les méthodes de l'équivalence locative et du coût d'utilisation sont essentiellement différentes à cause de l'imposition différente des personnes et des biens. De plus, si l'on tient compte de la possibilité de marchés du logement différents (entre

la location et la propriété), alors la tentative d'évaluer les coûts d'utilisation par l'équivalence locative est vouée à un échec d'estimation encore plus important. Bien que les coûts d'utilisation présentent des problèmes d'estimation plutôt complexes en ce qui concerne la définition des variables appropriées et leur collecte, la technique d'équivalence locative n'est pas par ailleurs suffisamment justifiée au point de vue théorique pour être utilisable comme mesure du coût de la propriété d'une maison.

I. Introduction

Inflation has been one of the most important economic problems in the last decade or so. Measurement of inflation is, therefore, an important economic issue not only for theoretical interest but also for its significant policy implications. Indices of inflation are used to make decisions on pricing, wage negotiations and social benefits calculations, to name a few examples. The so-called Consumer Price Index is the measure most frequently used to determine cost of living and the rate of inflation. As is well known, the CPI is a Laspeyres index that uses the base period consumption basket as the weights. It tends to overestimate cost of living, since substitution behaviour of consumers is not taken into account. Basically any cost-of-living index is a ''weighted average'' (or variant thereof) of ''relevant'' prices. Thus the two key issues are (1) what goods should be included and (2) what weights should be attached to each good. A number of interesting issues arise in an attempt to construct a meaningful index of cost of living.

In this paper, we investigate a particular issue: how to account for ''housing'' or more specifically, the ''shelter'' component of the cost of living. (We do not, therefore, consider ''utilities'' or ''household furnishing and operation'' components.) Consumer durable goods in general raise additional and unique problems in cost-of-living calculation. By definition the consumption of durables extends over more than one period. Thus purchase and consumption do not coincide. In this case, the observed market data at a point of time on purchase prices and quantities would have to be translated to account for the flow of services per period. Housing (shelter), among consumer durables, presents a more complex and serious case to deal with: partly because its lifetime is long and partly because its expenditure share is significant. Thus, estimates of housing cost have a large impact on the overall CPI.[1] However, the problems associated with durability are common among other consumer durables.

We focus on two approaches to the measurement of housing cost: the rental equivalence and the user cost approaches. The first asks how much the homeowner would charge someone such as himself to rent his house, while the second approach focuses on the opportunity cost of owning and using the house. These are conceptually equivalent, but may be different operationally. The rental equivalence approach is usually interpreted as a method of estimating potential rent by observing actual market rents. Unfortunately, observed market rents and user cost estimates seem to differ substantially (e.g. McFadyen and Hobart [1978] and Hendershott [1980]). Why are they so different? This large discrepancy is puzzling and unsettling. What do these measures represent? Which should be used?

In this paper, we consider the impact of taxes on user costs and market rents. The major features of the Canadian and U.S. tax systems that affect housing costs are highlighted in order to contrast the relevant variables which compose the two housing measures. We shall argue that tax considerations are important in choosing the appropriate housing measure.

The paper is organized as follows. In Section II, we discuss basic issues surrounding these two approaches. Section III presents a simple framework for analyzing the user cost and rental approaches. We then discuss the implications of the model on the CPI estimates.

II. Relevant Issues

The current treatment of housing cost in the CPI has been criticized by many both in Canada and the United States (see Gordon [1981] and Gillingham [1982] for a good summary). It appears that both Statistics Canada and the BLS are keenly aware of the inadequacy of the current methods as well as the specific problems associated with alternative approaches. Nevertheless, there are significant practical barriers to change. For example, in facing a new proposal for a user cost approach, the Price Committee of the BLS Labor Research Advisory Council rejected it on the ground that it would have introduced "an inappropriate hybrid condition into the historically well established and consistent price transaction basis for the constant market CPI".[2]

The user cost estimates have provided reasonable results in explaining consumer tenure decision (Hendershott and Shilling [1982], Boehm and McKenzie [1982], and Rosen and Rosen [1980]), while performing poorly in another (Muellbauer [1981]). Unfortunately the estimation of user cost is saddled with many serious problems: how to define and measure variables, how to treat capital gains (accrued or realized, *ex ante* or *ex post*), and how to account for the differential tax treatment of individuals and assets.

The rental equivalence approach seems to have gained favour among some economists (e.g. Gillingham, Rymes and Gordon) on either theoretical or operational grounds. In fact, the BLS has introduced the experimental rental equivalence measure in 1982 and is planning to switch completely to this measure in early 1983.[3] By using observed market rents, the collection of necessary data is much more simplified than that for the user cost calculation.

Gillingham [1980] advocates the rental equivalence approach as the better of the two alternatives. After calculating user cost indices for 1964-75, using different values for the expected rate of home price appreciation, cost of capital and the mortgage rate, he concludes that the resulting estimates are substantially different. This exercise, he argues, demonstrates "the sensitivity of the indices to alternative assumptions about the individual user cost components" and supports the position that "it is impossible to construct a valid user cost measure which is consistent with the information provided by rent market...". During the period chosen by Gillingham, significant upward but volatile changes in most goods prices as well as interest rates were observed. Therefore, it is not surprising that the calculated indices are very sensitive to the values chosen. We believe that the proper conclusion to make from this exercise is not that user costs are unreliable, but that appropriate selection and estimation of relevant variables is essential. Of those variables mentioned, estimates of housing price increase rates are clearly critical. More will be said on this point later.

On the other hand, concerning estimates of the opportunity cost of equity capital and the mortgage rate, one should be able to make better choices or at least be able to narrow the range on theoretical grounds without resorting to permutations of possibilities. Why use current or five-year moving averages for the mortgage rate? We are not necessarily

in disagreement with the first part of his conclusion that ‘‘the rental equivalence approach provides a simple, more direct measure...’’. Nevertheless it needs to be demonstrated first that the observed market rents do represent what is desired, i.e. ‘‘the cost of shelter service for homeowners’’. We shall illustrate that tax considerations alone are sufficient for rents to diverge from user costs. In that case, rents may be used only as a proxy (possibly a poor one) for the cost of housing service for owner-occupied homes.

Gordon [1981] is critical of the general method used to calculate homeownership cost for the American CPI. (The same can be said about the Canadian CPI.) His objections are: (1) the index is calculated as if all homeowners are subject to current mortgage rates; (2) the weight of homeownership cost is too high due to double counting of purchase price and mortgage payments and (3) the CPI index does not take the deductibility of interest payment (U.S. only) and capital gains into account.

Gordon is also pessimistic about the feasibility of estimating user cost due to the measurement problem. Noting that user costs tend to be much more volatile than market rents, he concludes that ‘‘If an economist’s approximation of how much a house should rent for does not behave at all like actual observed rents, then that ought to be telling him something.’’[4]

What does it tell us? It may be telling us that our approximation is poor. Or alternatively, it might be telling us that rents and user cost are not necessarily the same, especially when there are many institutional distortions in existence. The latter position is supported even more strongly if rental housing and owner-occupied homes do not belong to the same market.

In summary, there are three arguments frequently put forth in favour of the rental equivalence approach.

(1) The user cost estimates and market rent do not move together. In addition, cost estimates tend to be volatile, whereas rents have been relatively stable.

(2) The estimation of user cost presents serious difficulty in defining and estimating rele-

vant variables (e.g. the expected rate of inflation).

(3) The collection of data for user cost estimation is more expensive. BLS [1977: 13] states that ''collecting (housing) prices for local regions is expensive''. Why is it more expensive than collecting apple prices?

The BLS has been experimenting with a new approach to the CPI by placing greater emphasis on the rental equivalence index.[5] The new CPI index, CPI-X1, uses the CPI rent component as the shelter measure. Obviously, this requires collection of data on rents from a wider sample of rented homes than was done previously, if this index is to be viewed as the user cost of owner-occupied single family dwellings.

There may indeed be an advantage in using the rental data: the definition and collection of data is simpler and there is less room for disagreement. Perhaps this is advantageous from a political standpoint. On the other hand, one has to establish a sound theoretical justification for using a particular method. One should, in addition be aware of the possible errors made by not using the alternative. That the decision to switch to the rental equivalence approach by the BLS is based upon ''conceptual and operational adequacy'' (Gillingham [1982:14] strikes us as only a weak form of justification). In the next section, we will develop simple theoretical models for the user cost of owner-occupied homes and the rental cost for tenant-occupied homes. These models are designed to incorporate the major characteristics of tax treatments in the United States and Canada: In particular, we focus on the differential tax treatments of individuals and assets.

III. The Model

We assume that a typical family makes allocation decisions in such a way to maximize household utility. The family allocates their wealth over goods and leisure for the current period as well as over time. Thus the decision whether to purchase or rent a home, what type of housing to acquire (quality, size and location), and how long to keep the home, depends on the preferences of the family (the demographic composition would be important here), and its budget constraint.

Suppose we take a family which has made an optimal decision to purchase a home. What is its opportunity cost of using the house for accommodation? Let us first assume a "perfect" world in which neither transactions cost nor taxes exist. In addition, we assume perfect capital, rental and resale markets. The user cost (Jorgenson [1971]) in this world would be:

$$V_t = (r_t + \delta - \dot{p})\, P_h + M_t \,, \tag{1}$$

where V_t is the user cost, r_t is the "appropriate" rate of return on the resources used to purchase the home, δ is the rate of depreciation, $\dot{p}$ is the (expected) growth rate in housing prices, P_h is the current price, and M_t is the cost of maintenance, repairs and insurance at time t. All variables are in nominal terms. Equation (1) accounts for the opportunity cost of financial resources tied up in the house, changes in physical quantities, property maintenance and changes in market value. This equation is used as the basic model.

In a world without taxes, transactions costs or uncertainty, the rental cost (charged to the tenant) and the user cost (implicitly charged to the homeowner) will be equal to V_t in long run equilibrium (see Dougherty and Van Order). We now modify (1) to incorporate Canadian and U.S. taxes. In the United States, all interest payments and property taxes are deductible from taxable income. In both countries, imputed rents are not taxed whereas rents received by landlords are taxed.

The Case of United States

For the U.S. owner-occupied homeowner, the user cost will be:

$$V_t = [\,(1-\tau_i)r_t + \delta - \dot{p}]P_h + M_t + (1 - \tau_i)PT_t \,, \tag{2}$$

where τ_i is the marginal tax rate of the homeowner i and PT_t denotes the property taxes at time t. It is clear from (2), that the user cost differs for two families with identical houses, if they are in different tax brackets. In fact, the deductibility of interest payments and property taxes is more advantageous, the higher the family's marginal tax rate. That is,

the after-tax user cost declines with the family's taxable income. If the opportunity cost of the downpayment is not the same as the mortgage rate, then the interest term becomes a weighted average of the mortgage rate, r_m, and the market opportunity cost, r_e. In this case,

$$V_t = \left\{ [\alpha\, r_e + (1 - \alpha) r_m] (1 - \tau_i) + \delta - \dot{p} \right\} P_h + M_t + (1 - \tau_i) PT_t , \qquad (2')$$

where α is the downpayment as a fraction of the purchase price. We have so far assumed that the homeowner costlessly purchased and sold the house (to himself) at every period. In such a case, the appropriate opportunity cost of capital is reflected by current interest rates. On the other hand, many American homeowners are subject to a fixed mortgage rate. (In fact, many are benefiting from relatively low interest rates from the 60s and most of the 70s.) In this case, the basic equation is modified further:

$$V'_t = [r_t (1 - \tau_i) + \delta - \dot{p}] P_h - (1 - \alpha_0)(r_t - r_{t0})(1 - \tau_i) P_{h0} + M_t + (1 - \tau_i) PT_t ,$$

$$= V_t - (1 - \alpha_0)(r_t - r_{t0})(1 - \tau_i) P_{h0} , \qquad (2'')$$

where 'P_{h0}' is the initial purchase price, α_0 is the proportion of the purchase price paid in cash as a downpayment and r_{t0} is the fixed mortgage rate at the time of purchase. Thus when r_{t0} is less than r_t, the user cost is reduced by the amount represented by the second term in (2''). We assume that the owner carries the original rate, when he sells the house back to himself at every period. Most mortgage financing until recently was at a fixed rate for 25 to 30 years; thus the equation (2'') may be a more accurate reflection of reality. It is easy to see that (2) and (2'') are the same when $r_{t0} = r_t$.

The Case of Canada

Mortgage (or any) interest payments are not deductible in Canada. It may appear that we can apply equation (1) to Canadian homeowners, however, a modification becomes necessary due to the fact that the returns from alternative investment opportunities are generally taxable. Facing the choice between putting up a larger cash downpayment (or

repaying the mortgage loan) and buying, say, government bonds, the homeowner should rationally select the former alternative. He realizes that the rate of return to cash downpayment is the full amount of the mortgage rate, whereas that on government bonds is $(1 - \tau_i)r_t$.

Thus for a homeowner with 100% of the purchase price financed externally (assuming again that a costless transaction takes place at every period):

$$V_t = (r_t + \delta - \dot{p})\, P_h + M_t + PT_t\,, \qquad (3)$$

whereas for a homeowner who paid cash for the entire purchase price:

$$V_t = [\, r_t\,(1 - \tau_i) + \delta - \dot{p}]\, P_h + M_t + PT_t\,. \qquad (3')$$

If α% of the purchase price is paid by cash and the rest financed, then:

$$V_t = [\, r_t(1 - \alpha\tau_i) + \delta - \dot{p}\,]\, P_h + M_t + PT_t\,. \qquad (4)$$

In general, we have:

$$V_t = [r_e(1-\tau_i)\,\alpha + r_m(1-\alpha) + \delta - \dot{p}]P_h + M_t + PT_t\,. \qquad (4')$$

Since the best alternative for mortgagors is early repayment, $r_e = r_m$. In this case, (4) = (4'). It is clear that the Canadian homeowners have an incentive to repay their mortgages as soon as possible.

The Case of Rental Housing

In order to contrast homeownership cost with cost of renting, we investigate rental markets. First we focus on the extreme case in which the owners of the rental units are not able to avoid taxes on appreciation in the house price (either in the form of capital gain or ordinary profit). In this case, all of their cash inflows and outflows are subject to tax: The inflows are taxed and the outflows are deductible. In addition, at the time

of sale, any appreciation is taxed. Thus, in long run equilibrium, we expect that rents will just cover the opportunity cost of the investment. Therefore, rent will be equal to the sum of all costs involved, i.e.

$$R_t = (r_t + \delta - \dot{p}_R)P_h + M_t + PT_t , \qquad (5)$$

where r_t is the risk-adjusted rate of return, $\dot{p}_R$ is the expected rate of appreciation in the rental housing market. $\dot{p}_R$ is equal to $\dot{p}$, if the rental and ownership market are the same. (Of course, for any particular house, the rate of appreciation may be the same whether it is rented or owner-occupied, *ceteris paribus*.) The after-tax relationship is equivalent to the before-tax relationship, if all terms are subject to the same tax rate.

Landlords are typically allowed to deduct depreciation costs based on historical rather than replacement cost. During the period of inflation, this arrangement is disadvantageous to the rental property owner, since the deductions are made smaller in real terms. In this case, we expect the rental cost to increase to cover a higher opportunity cost.

For those individual investors who manage to avoid or reduce capital gains tax by exchanging properties or deferring sales, the after-tax cost of holding a rental property is lowered further. In this case,

$$R_t = [r_t + \delta - \dot{p}(1 - \tau_c)/(1 - \tau_i)]P_h + M_t + PT_t , \qquad (5')$$

where τ_c is the tax rate on capital gain and τ_i is the marginal tax rate as before.

IV. Observations and Discussion

(1) The user cost equations above clearly illustrate that the tax laws can be a significant factor in affecting user cost. Differences in the tax treatment between Canada and the U.S. residents lead to different calculations. In both cases, the tax bracket of a particular homeowner is relevant. That is the higher the marginal tax rate, the lower the after-tax user cost. (See Villani [1982] for a detailed computation of how marginal tax rates and

inflation rates affect cost of capital.) Therefore, we expect both the income effect and substitution effects at work as the family income increases.[7]

(2) The Canadian homeowners can lower the user cost further by increasing their cash downpayment or repaying principal on their mortgage. On the other hand, there is no such built-in bias in the U.S. In fact, when r_{t0} is less than r_t, the U.S. homeowners have no incentive to repay their mortgage any faster than they are required. It is not surprising that the percentage of owner-occupied units with mortgages has increased from 45.3% to 64% from 1940 to 1979.

(3) The rental equations illustrate the difference between user cost and rent. This difference is due mainly to three factors: (i) rental income is taxable whereas imputed rent is not; (ii) depreciation, maintenance, etc. are tax deductible for owners of rental units only; (iii) landlords are taxed on appreciation in housing prices, while homeowners are exempt.

Our simple framework implies that homeowners in both countries are subsidized by the tax treatment compared with rental owners and tenants. Consequently, the price of homeownership is lowered relative to that of renting. The tenure decision, we would expect, should be influenced by this tax break. Rosen and Rosen [1980] estimated that about one-fourth of the increase in homeownership in the United States since 1945 is attributed to the favourable tax treatment of homeownership.

The above observations lead to the conclusion that the rental equivalence approach (which may have an operational advantage over the user cost approach) presents a serious problem. It is not obvious how observed market rent is related to the user cost of owner-occupied homes. Two points are in order. The first is that landlords and homeowners are treated prejudicially by the tax system. Thus, even if we have two housing markets, one rental and the other owner-occupied, but otherwise identical, we would not necessarily expect the rent charged by the landlords to be equal to the imputed rent which the homeowner charges (to himself) in long run equilibrium. The occasional homeowner landlords who rent their homes during their sabbatical years may attempt to charge V_t, but it is not clear to us if they are in any better position to calculate V_t than those who

are in charge of estimating the CPI shelter costs. Thus the reliability of observed market rent depends upon whether there exists a well established and efficient rental market for the types of housing in question. In the event we obtain reliable rental figures for the owner-occupied equivalent homes, we can adjust these to estimate the imputed user cost.

Second, it is possible that the housing market is segmented into two markets: one for rental and the other for ownership. This may be due to differential tax treatment, institutional factors such as zoning and most importantly, to the type of housing typically demanded and supplied in the two markets. Demographic characteristics of consumers tend to be quite different.[8] Presumably, it may be necessary to formulate a general equilibrium model of the residential housing market in order to investigate the relationships between rental and homeownership markets.

User cost estimates have been criticized for their volatility, due basically to volatile estimates of the rate of (expected) appreciation. Various methods have been used to estimate this variable. Examples of methods used are: actual price increases (McFadyen and Hobart [1978]), distributed lags of past changes (Rosen and Rosen [1980] and Gillingham [1980]), and distributed lag model of weighted averages of general inflation rate and housing price increases (Hendershott and Shilling [1982]). It is not clear what should be used *a priori*; we do not have a good understanding of how expectations are formed. Using past price changes over a few periods tended to result in fairly volatile estimates and therefore, we should expect the same thing in estimating rental user cost (Hendershott [1980]). Of course, observed market rents are much more stable. Does this imply, according to Gordon's lesson, that the rental user cost estimates are incorrect? It could be that rents are determined by forces other than what we accounted for. Our model then is mis-specified. Alternatively, even if the model is correctly specified, we may not have used proper estimates of the variables (errors in measurement). For example, the rate of expected price appreciation for landlords as well as homeowners may be much more stable than what researchers have hypothesized in their studies. The *ex post* and *ex ante* appreciation rates may be quite different.

The use of first differences from previous periods has produced estimates of negative user cost. Negative prices are rejected because they imply infinite consumption. However,

if the goal of our exercise is to measure *ex post* "welfare" of households such as in national accounting, the *ex post* improvement in the wealth of some consumers due to (realized or unrealized) capital gain on their homes should be included. After all, capital gains can always be realized and used for consumption of other goods and services. (Of course, the rate of appreciation of all or most goods must be less than capital gains in order to observe this effect.) On the other hand, in making policy decisions, for which the CPI index may be an important component, negative user costs are not desirable, even though the CPI index is an *ex post* accounting measure. The CPI index is often used for future adjustments. Negative indices, even for a subindex such as for shelter, is neither consistent with our theory (*ex ante*) or evidence. Speculative activities seem to have been stimulated during periods of high price appreciation, but purchases were certainly not infinite. (Of course, supply constraint prevents this.)

Other factors, as well, may be significant. We have ignored transactions costs, information costs and other market imperfections: all of these tend to reduce the speed of adjustment by the consumers to changing environments. In addition, the purchase of durables, especially housing, tends to be lumpy.

There may be a reason to expect user cost to be positive most of the time. The rate of appreciation, even in terms of expectations, may not be independent of other variables. A higher rate will, *ceteris paribus*, reduce the user cost in our formulation. However, higher rates may be associated with higher values for other variables. For example, a study by Boehm and McKenzie [1982] claims that an increase in expected appreciation by 1% reduces demand for housing by 3%. In this case, the appreciation (and the tax advantage due to higher marginal tax rates) is outweighed by the increase in the interest rate.

In summary, we wish to point out again that the seemingly more desirable approach to the shelter index via the rental equivalence approach also presents some serious problems both in terms of theoretical justification and operational feasibility. This is due to the inherent deviation that exists between the user cost of homeowners and the rental cost of landlords. We have shown that this is caused partially by differential tax treatments of individuals and assets, and partially by the segmented housing markets. Thus, if one wishes to account for this deviation, it is not a simple matter of applying the rental

equivalence approach. We are interested in measuring the rate of changes in consumers' housing cost ''effectively and appropriately''. The rental equivalence is certainly a viable approach to this purpose, however, whether or not that is the best approach is yet to be established.

Footnotes

[1] How housing is treated in the CPI is of critical importance in the overall CPI computation, since the weight is very large. For example, the weights attached to the ''shelter'' component of the CPI are as follows:

CANADA

1978 Basket

	Weights (total = 100)
Shelter	20.66
Rented Accommodation	7.18
Owned Accommodation	12.59
Other Accommodation	.63

Statistics Canada [1982: 138]

USA 1977 (All Urban Consumers)

Shelter	29.181
Rent, Residential	5.624
Other Rental Costs	.711
Homeownership	22.846
Home Purchase	9.967
Financing, Taxes and Insurance	9.211
Maintenance and Repairs	3.688

BLS[1978: 8]

[2] BLS [1977: 14]

[3] Consequently, the weights of the homeownership and shelter components are reduced approximately by one-half and one-third respectively for 1980. See Gillingham [1982: 12] for more details.

[4] Gordon [1981: 126].

[5] See Norwood [1981] and Gillingham [1982].

[6] It is a simple arithmetical exercise to see the magnitude of the user cost difference. For example, if we have $r_t = .10$, $p_h = \$100{,}000$, $PT = \$1{,}000$ and $\delta = p = 0$, then $V_t = \$9{,}300$ for a family of $\tau_i = .2$ and $V_t = \$6{,}000$ for another with $\tau_i = .5$.

[7] Rates of homeownership in Canada for 1978 indeed increases from 26.5% to 86.1% as the family income increases from less than \$6,000 to more than \$35,000.

[8] The "typical" dwelling in the rental and the owner-occupied markets are as follows:

	Renter-Occupied	Owner-Occupied
No. of Bathrooms	1.1	1.58
No. of Rooms	4.1	6.16
No. of Bedrooms	1.81	3.0
Age	17.65	17.44
(Subjective) Excellence Scale	2.15	1.59
Lengh of Tenure	4.91	12.08

Follain [1978: 11]

The studies on heldonic price on housing tended to separate the rental from the ownership market and have obtained substantially different coefficient estimates on characteristics. See Saccomanno [1979].

References

Boehm, Thomas P. and McKenzie, Joseph A. [1982], "Inflation, Taxes and Demand for Housing", *AREUEA Journal*, Spring, 10(1): pp.25-38.

Bureau of Labor Statistics [1977], The Consumer Price Index: Concepts and Content Over the Years, Report 517.

_________ [1978], The Consumer Price Index: Revision – 1978.

Dougherty, Ann and Van Order, Robert [1982], "Inflation, Housing Costs and the Consumer Price Index", *American Economic Review*, March, 72(1): pp.154-64.

Follain, James R. Jr. [1978], "Cross-Sectional Indexes of the Price of Housing", Draft.

Gillingham, Robert [1980], "Estimating the User Cost of Owner-Occupied Housing", *Monthly Labor Review*, February, 103(2): pp.31-5.

_________ and Lane, Walter [1982], "Changing the Treatment of Homeownership in the CPI", *Monthly Labor Review*, June 105(6): pp.9-14.

Gordon, Robert J.[1981], "The Consumer Price Index: Measuring Inflation and Causing It", *The Public Interest*, Spring, 63: pp.112-34.

Greenlees, John S. [1982], "An Empirical Evaluation of the CPI Home Purchase Index, 1973-78", *AREUEA Journal*, Spring, 10(1): pp.1-24.

Hendershott, Patrick H. [1980], "Real User Costs and the Demand for Single-Family Housing", *Brookings Papers on Economic Activity*, 2: pp.401-52.

_________ and Shilling, James D. [1982], "The Economics of Tenure Choice", in *Research in Real Estate: A Research Manual*, C.F. Sirmans (ed.), Vol. 1, JAI Press: Greenwich, Connecticut.

Jorgenson, Dale W. [1971], "Econometric Studies of Investment Behavior: A Survey", *Journal of Economic Literature*, December, 9(4): pp.1111-47.

McFadyen, Stuart and Hobart, Robert [1978], "An Alternative Measurement of Housing Costs and the Consumer Price Index", *Canadian Journal of Economics*, February, 11: pp.105-12.

Muellbauer, John [1981], "Testing Neoclassical Models of the Demand for Consumer Durables", in *Essays in the Theory and Measurement of Consumer Behavior*, Angus Deaton (ed.), Cambridge University Press: Cambridge.

Norwood, Janet [1981], "Two Consumer Price Index Issues: Weighting and Homeownership", *Monthly Labor Review*, March, 104(3): pp.58-9.

Rosen, Harvey and Rosen, Kenneth T. [1980], "Federal Taxes and Homeownership: Evidence from Time Series", *Journal of Political Economy*, February, 88(1): pp.59-75.

Rymes, T.K. [1979], "The Treatment of 'Homeownership' in the CPI", *Review of Income and Wealth*, December, 25: pp.393-412.

Saccomanno, F.F. [1979], "A Hedonic Price Index for Housing in Metropolitan Toronto: 1961 and 1971", Research Paper No. 103, Univerity of Toronto.

Statistics Canada [1978], Family Expenditure in Canada, 62-550.

__________ [1982], The Consumer Price Index Reference Paper: Concepts and Procedure.

Villani, Kevin [1982], "The Tax Subsidy to Housing in an Inflationary Environment: Implications for After-tax Housing Costs" in *Research in Real Estate*, Vol. 1, C.F. Sirmans (ed.), JAI Press: Greenwich, Connecticut.

THE INCORPORATION OF DIRECT TAXES INTO A CONSUMER PRICE INDEX

Robert Gillingham and
*John S. Greenlees**
Office of Prices and Living Conditions
Bureau of Labor Statistics

SUMMARY

The U.S. Consumer Price Index (CPI) measures the change over time in the cost of a fixed market basket of goods and services. It can be interpreted as a fixed-weight approximation to a conditional cost-of-living index, where (1) the cost of living is defined as the minimum expenditure necessary to achieve a particular level of satisfaction and (2) the cost is defined to be conditional on all of the determinants of the level of satisfaction except current quantities of market goods and services. The importance of this interpretation stems from the widespread belief that the measurement objective of the cost-of-living index is appropriate for many of the uses to which the CPI is put, including wage escalation and deflation of income. What is often ignored, however, is that many of the value aggregates being either escalated or deflated are components of before-tax income. Unfortunately, escalation or deflation of before-tax income by the CPI need not result in a value aggregate of equivalent real purchasing power. In this paper we will define and estimate an index which incorporates direct as well as indirect taxes into a CPI. Following U.K. terminology, we refer to this index as a Tax and Price Index (TPI). The methodology we use is not a substitute for or correction of the current CPI methodology. Rather it approximates an alternative measurement objective – the before-tax income necessary to achieve a given level of satisfaction – which may be more appropriate for at least some of the uses to which the CPI is put.

The remainder of the paper is organized as follows. In Section 2 we describe in some detail the index which we calculate and its relationship to the cost-of-living framework.

In Section 3 we describe the data and algorithms used in the estimation process. An important aspect of our approach is that we estimate indexes at the individual household level. The benefits of this approach become apparent in the next two sections. In Section 4 we present estimated aggregate indexes for both the entire population on which we focus and selected partitionings of this population for the period from 1967 to 1981. In Section 5 we analyze at the micro level the relationship between tax burdens and selected household characteristics and how this relationship changed between 1967 and 1981. Finally, in Section 6 we summarize our results and outline plans for future research.

RÉSUMÉ

L'indice des prix à la consommation (IPC) des États-Unis mesure la variation dans le temps du coût d'un ensemble de biens et de services offerts sur le marché. On peut l'interpréter comme une approximation à pondération fixe d'un indice conditionnel du coût de la vie, où (1) le coût de la vie est défini comme la dépense minimale nécessaire pour obtenir un niveau particulier de satisfaction et (2) le coût est par définition conditionné par tous les déterminants du niveau de satisfaction, sauf les quantités courantes de biens et de services marchands. L'importance de cette interprétation vient de la croyance répandue voulant que l'objectif de mesure associé à l'indice du coût de la vie convienne pour nombre des utilisations auxquelles se prête l'IPC, y compris l'ajustement des salaires et la déflation du revenu. On néglige toutefois souvent le fait que plusieurs des agrégats faisant l'objet d'un ajustement ou d'une déflation sont des éléments du revenu avant impôt. Malheureusement, l'ajustement ou la déflation d'un revenu avant impôt au moyen de l'IPC ne donne pas nécessairement un agrégat de pouvoir d'achat réel équivalent. Dans ce document, nous définirons et estimerons un indice qui incorpore les impôts tant directs qu'indirects à un IPC. Reprenant en cela la terminologie britannique, nous parlerons d'Indice d'impôt et de prix (IIP). La méthodologie que nous utilisons ne vise pas à remplacer ou à corriger la méthodologie actuelle de l'IPC. Elle vise plutôt un autre objectif de mesure – le revenu avant impôt nécessaire pour obtenir un niveau donné de satisfaction – qui convient peut-être mieux, au moins pour certaines des utilisations auxquelles se prête l'IPC.

Ce document est organisé comme suit. Le chapitre 2 décrit de façon assez détaillée l'indice calculé et son rapport avec le cadre du coût de la vie. Au chapitre 3, nous décrivons

les données et les algorithmes utilisés pour l'estimation. Un aspect important de notre méthode est que nous estimons les indices au niveau du ménage particulier. Les avantages de cette façon de procéder apparaîtront dans les deux chapitres suivants. Le chapitre 4 présente les indices agrégatifs estimés tant pour la population totale que pour des partitions choisies de cette population, pour la période allant de 1967 à 1981. Au chapitre 5, nous analysons au niveau microéconomique le rapport entre le fardeau fiscal et des caractéristiques choisies des ménages, tout en examinant la façon dont ce rapport a évolué entre 1967 et 1981. Au chapitre 6, enfin, nous résumons nos résultats et présentons des plans de recherches futures.

1. Introduction

The U.S. Consumer Price Index (CPI) measures the change over time in the cost of a fixed market basket of goods and services. It can be interpreted as a fixed-weight approximation to a conditional cost-of-living index, where (1) the cost of living is defined as the minimum expenditure necessary to achieve a particular level of satisfaction and (2) the cost is defined to be conditional on all of the determinants of the level of satisfaction except current quantities of market goods and services.[1] The importance of this interpretation stems from the widespread belief that the measurement objective of the cost-of-living index is appropriate for many of the uses to which the CPI is put, including wage escalation and deflation of income. What is often ignored, however, is that many of the value aggregates being either escalated or deflated are components of before-tax income. Unfortunately, escalation or deflation of before-tax income by the CPI need not result in a value aggregate of equivalent real purchasing power. In this paper we will define and estimate an index which incorporates direct as well as indirect taxes into a CPI. Following U.K. terminology, we refer to this index as a Tax and Price Index (TPI).[2] The methodology we use is not a substitute for or correction of the current CPI methodology. Rather it approximates an alternative measurement objective – the before-tax income necessary to achieve a given level of satisfaction – which may be more appropriate for at least some of the uses to which the CPI is put.

The remainder of the paper is organized as follows. In Section 2 we describe in some

detail the index which we calculate and its relationship to the cost-of-living framework. In Section 3 we describe the data and algorithms used in the estimation process. An important aspect of our approach is that we estimate indexes at the individual household level. The benefits of this approach become apparent in the next two sections. In Section 4 we present estimated aggregate indexes for both the entire population on which we focus and selected partitionings of this population for the period from 1967 to 1981. In Section 5 we analyze at the micro level the relationship between tax burdens and selected household characteristics and how this relationship changed between 1967 and 1981. Finally, in Section 6 we summarize our results and outline plans for future research.

2. Specification of a Tax and Price Index

The U.S. CPI, as it is currently compiled, measures changes in the **expenditure** necessary to consume a fixed set of goods and services.[3] The CPI is best interpreted within the conditional cost-of-living index framework introduced by Pollak [1975]. It approximates a cost-of-living index which focuses on the current consumption of market goods, where market goods are defined as those goods to which a user charge is attached, regardless of the supplier. Within a multi-period framework, it is conditional on the future consumption of market goods, current and future consumption of public goods and leisure, and current and future environmental conditions. That is to say, in comparing alternative price vectors for current market goods, its cost-of-living interpretation is restricted to alternatives in which all other variables affecting satisfaction levels are assumed fixed.[4] The restricted focus of the CPI stems from operational considerations rather than a theoretical desire to limit its scope. Substantial work, some of which is the subject of this conference, is aimed at broadening this coverage. Regardless of the exact coverage, however, there is an obvious alternative to the expenditure focus of the current CPI which lies clearly within its cost-of-living orientation.

For many of the uses to which the CPI is put, it is perhaps more reasonable to measure changes in the income, before taxes, a consumer must receive to achieve a given level of satisfaction. Pollak [1972] refers to this index as an income-based cost-of-living index (ICOL) as opposed to an expenditure-based cost-of-living index (ECOL). In many of the situations in which the CPI is used to escalate all or part of a consumer's income, the focus

of escalation is a component of before-tax income – for example, money wages or retirement income. There are two obvious alternative rules which could be followed in determining the rate of escalation. The first is to escalate the component of income at the rate at which total income must change in order to maintain economic well-being. The second is to escalate the component by an amount which, when combined with the rate at which other components are changing, will yield the total change necessary to maintain economic well-being. The first objective will be met by an ICOL, while the second requires adjustment of an ICOL to account for differential rates of change in the remaining income components. For both cases an ICOL, or an approximation to an ICOL, is necessary, and in neither case will escalation by the current, expenditure-based CPI be appropriate.

In this paper we will estimate an approximation to an ICOL analogous to the CPI approximation to an ECOL. As in the CPI, the basis for the index will be a fixed-weight consumption pattern. In addition to the consumption components included in the CPI, we also include charitable contributions. We ignore savings.[5] The before-tax income which we measure will be the minimum necessary to yield an after-tax income equal to the expenditure required to purchase the fixed set of consumption goods. We calculate three ''tax'' components covering federal and state tax and Social Security contributions, respectively. As in the CPI, possibilities for substitution in response to changes in relative prices will be ignored. We treat Social Security contributions as a tax because the relationship between changes in real contributions and changes in real expected benefits is sufficiently tenuous to make this a reasonable first approximation. The index we calculate is essentially a CPI with tax components added, and, as mentioned in the introduction, we refer to this index as a Tax and Price Index (TPI).

We emphasize that construction of an ICOL is a conceptually distinct exercise from that of extending the scope of the cost-of-living index to include public or environmental goods. Recognition of these factors in a cost-of-living index is independent of the treatment of taxes and is beyond the scope of this paper.

To calculate the expenditure portion of a TPI, we need know nothing about a consumer unit other than its consumption pattern. To calculate the tax components of the TPI, however, it is necessary to know a number of economic and demographic characteristics

of the consumer unit and to establish a number of conventions. A household's tax liability depends upon the following, household-specific factors: (1) household composition, (2) income source and (3) consumption pattern. Much of the impact of these factors is fairly straightforward. For example, household composition affects federal and state tax liability through its impact on, *inter alia,* filing status, number and type of exemptions, and eligibility for special programs. Because of special rules concerning the tax treatment of, e.g., medical care and sales tax, a household's tax liability also depends on its consumption pattern. Social Security contributions depend upon whether the source of income is wages and salaries or self-employment. In addition, income source can determine eligibility for special federal tax credits. Each of these examples is essentially straightforward, and our ability to incorporate them into the calculation of the index is limited only by the practical problems related to compiling and implementing all of the rules for tax computation.

The interaction between a household's consumption pattern and income source can affect the calculation of TPI in a much more complex fashion, requiring the establishment of a number of conventions. Ownership of consumer durables by a household implies that a portion of the household's income is in kind and that the amount of in-kind income is equal to consumption of the services of the durable goods. With the exception of housing, we adopt the convention that expenditures on the stock of the durable good are a reasonable approximation of the value of the services consumed. This enables us to focus on a money transaction for these durables in the consumption index and abstract from the treatment of in-kind income in the tax components.[6] For shelter, however, we adopt an explicit flow-of-services approach in the consumption index, and this requires explicit consideration of the treatment of in-kind income in calculating the tax components of the TPI.

The preferential tax treatment of owner-occupants in the U.S. federal and state tax codes stems from the fact that the implicit rent which they receive is not taxed. We build this fact into our index by assuming that the amount of in-kind income for homeowners is identical to the value of shelter services consumed and that both of these amounts are appropriately indexed with the CPI rent index. Consequently, the TPI for homeowners measures the total money and in-kind income necessary to yield an after-tax income equal to the explicit cost of non-shelter consumption plus the implicit cost of shelter.[7]

3. Construction of Indexes

3.1 Data

In order to compute our tax and price indexes for the period 1967-81, we required three major categories of data. These were (1) federal individual income tax and Social Security tax regulations, (2) state personal income and sales tax rules and (3) household-specific consumption cost series. Below we briefly describe our data sources and the procedures used to compute such parameters as total consumption, filing status and itemized deductions.

Federal Tax Data. Information on individual income tax brackets, marginal rates, exemptions, deductions, and credits for the years 1967 through 1979 was drawn from *Individual Income Tax Returns*, an annual series produced by the Internal Revenue Service (IRS). Each volume contains copies of and instructions for the 1040, 1040A and related individual return forms, along with discussions of relevant regulations and amendments. Since 1980 and 1981 volumes of the series are not yet available, for these tax years we referred to the IRS's guide to individuals, *Your Federal Income Tax*.

The essential structure of the federal tax system has remained unchanged since 1967. The tax due is computed by applying a system of increasing marginal rates to ''taxable income''. Tax credits may then be applied to reduce the tax bills of households in certain categories.

The first step in deriving taxable income from gross income is the computation of ''adjusted gross income'' (AGI). Subtractions from gross income include partial exclusions of specific income categories such as interest, as well as deductions for certain expenditures such as moving costs, employee business expenses and contributions to retirement funds. For the purposes of this paper we assume that all income derives from wages and salaries or from self-employment, and we exclude non-housing investment and business expenditures from consideration. We therefore treat gross income and AGI as equivalent concepts with the important exception of implicit rent, which we discuss in a later section.

Taxable income equals AGI less the value of exemptions and deductions. An exemption is given for each taxpayer, spouse and dependent, with additional exemptions for each taxpayer or spouse who is blind or aged 65 or older. (The elderly exemptions are reflected in our computed tax expenditures; however, we cannot identify blind individuals in our sample of households.) The most important categories of deductible expenses are medical care costs, state and local taxes, interest paid, and charitable contributions. A ''standard deduction'' is also available for taxpayers who do not choose to itemize or whose deductible expenses are relatively low.[8]

Rule changes during our study period have, for the most part, reduced the federal tax liability on a given nominal level of gross income. In 1967, the value of an exemption was $600. The standard deduction was set at 10% of AGI, with a maximum of $1,000 and a minimum of $200 plus $100 per allowable exemption. Tax rates for single taxpayers increased from 14% on the first $500 of taxable income to 70% on income over $100,000. The schedule for married couples was such that their tax was twice that which would be paid by a single person with one-half the taxable income – i.e., each of the 25 tax brackets was twice as wide as in the singles schedule.[9]

In 1968 a special 10% tax surchage was introduced and was in effect for the last nine months of the year. The rules for medical care deductions were also modified. Taxpayers and spouses aged 65 and over were no longer excluded from the rule that only medical care costs exceeding 3% of AGI could be deducted. At the same time, health insurance expenditures up to $150 were made exempt from the income limitation.

The 10% surcharge was in effect throughout 1969 and for three months in 1970. Other changes in 1970 were the increase in exemption value to $625 and the replacement of the minimum standard deduction by a relatively complicated formula based on AGI as well as on the number of exemptions.

In 1971 effective tax rates were reduced in several ways. For the first time in four years, there was no surcharge. The standard deduction percentage was increased to 13%, with a flat minimum of $1,050 and a maximum of $1,500. The value of an exemption was raised again, to $675. The rate schedule for single taxpayers was lowered, so that an individual

with a taxable income of Y now payed less than half the tax paid by a married couple earning 2Y. Finally, a maximum tax rate of 60% on ''earned income'' (the only form of income we deal with here) was introduced.

The rate schedules were modified twice more during our study period. In 1979 the brackets for both single and married taxpayers were widened, the total number of brackets being reduced from 25 to 16. Although the minimum and maximum rates remained at 14% and 70%, the average rate for any given income fell. In 1981, tax rates in every bracket were multiplied by a factor of .9875 to reflect the 5% tax cut in effect for the last three months of the year. The maximum tax on earned income was lowered in 1972, from 60% to 50%.

The value of an exemption continued to increase, to $750 in 1972 and $1,000 in 1979. The standard deduction also was liberalized in several stages. In 1975 separate minimums and maximums were established for singles and joint filers. In 1977 the minimums and the percentage deductions were deleted. In 1981 the now flat standard deduction stood at $2,300 for singles and $3,400 for married couples.

Our indexes reflect three federal tax credit programs which were introduced during the 1970s. The first is the General Tax Credit, which began in 1975 with the Personal Exemption Credit (PEC). Under the PEC, taxpayers could subtract $30 from their tax liability for each exemption (excluding the elderly and blind exemptions). This credit was increased to $35 in 1976 and extended to all exemptions in 1977, and beginning in 1976 taxpayers were allowed to take the greater of the PEC and the Taxable Income Credit, which equalled 2% of taxable income up to a limit of $180. These credits were dropped after 1978.

The Credit for the Elderly was introduced in 1976, extending previous tax credit programs for recipients of retirement income. The amount of the credit depended on filing status, AGI, and (for married couples) whether both taxpayer and spouse were more than 64 years old.

Finally, the Earned Income Credit (EIC) was instituted in 1975 to reduce the liability of low-income households. The program was liberalized in 1979, and as of 1981 the credit equalled 10% of AGI up to $5,000, less 12.5% of AGI above $6,000. Thus, the maximum

credit was $500, and the credit was zero for incomes above $10,000. One qualification requirement is that the household contain a dependent child who is aged 19 or under, disabled or a full-time student. The EIC is also unique among credit programs in that it is not limited by the pre-credit tax liability -- i.e., it can result in a negative total federal tax.

The second major component of the federal direct taxation system consists of contributions to the Social Security (FICA) retirement, disability and health insurance systems. These contributions are a constant proportion of earnings up to a ceiling level. Both the tax rate and the income ceiling have increased rapidly. In 1967, wage and salary workers contributed 4.4% of earnings below $6,600. By 1981 the tax rate and ceiling had reached 6.65% and $29,700, respectively. Self-employed individuals are subject to the same ceiling but contribute at a higher percentage rate, which rose from 6.4 in 1967 to 9.3 in 1981.

State Tax Data. We obtained data on state income and sales taxes from two Commerce Clearing House publications, the *State Tax Handbook* and *State Tax Guide*. In this section we discuss the most important variations among the state systems, and point out those aspects which we were forced to ignore due to resource constraints or the lack of available information.

The *State Tax Handbooks*, published annually, provided the tax brackets and marginal rates for each state and year, as well as the exemptions or credits given for taxpayer, spouse and dependents. From the current *State Tax Guide* we obtained information on rules for itemized deductions, elderly exemptions or credits, sales taxes and other details.

Of the 50 states and the District of Columbia, we excluded Alaska and Hawaii on the basis that national price series could not be treated as representative, and two other states, Montana and Wyoming, which were not included in our household sample. The remaining 47 jurisdictions included 16 which had no income tax at the beginning of 1967.[10] Part of the reason for the sharp increase in state taxes over our study period is that several states with large populations introduced income taxes between 1967 and 1981. These included Michigan in late 1967, Illinois in 1969, Pennsylvania in 1971 and Ohio in 1972.

Most states have progressive marginal tax rate systems similar to that of the federal

government. A few states have a single tax rate, while three compute tax liability as a percentage of the federal liability rather than by a rate schedule applied to income. Recently, several states have begun to index their tax brackets to a measure of the inflation rate.

Like the IRS, state systems generally give extra exemptions or credits to taxpayers, spouses or dependents aged 65 or over. Our computations assume that the current form of the elderly preference, as shown in the *State Tax Guide*, held throughout the 1967-81 period. Most commonly, states grant the same number of exemptions as are computed on the federal 1040 and 1040A forms.

We have no historical series on standard deductions at the state level. These tend to differ from the federal because state and local taxes are not deductible on state returns, while federal taxes sometimes are. We therefore must assume in our computations that all taxpayers itemize deductions on state returns. We do explicitly recognize differences among states in the deductibility of federal income tax, as well as the fact that some states do not allow any deductions from AGI in computing taxable income.

We leave for future research any recognition of county and city income taxes,[11] as well as the commuter taxes imposed by a number of jurisdictions. We also are forced to ignore variations in the deductibility of interest, energy-saving investments, medical care costs and other expenditures. With the exceptions noted above, we assume that the same categories of expenditure are deductible on state returns as on federal. Finally, we do not deal with special allowances or credits for low-income households, renters and other target groups.

We do not have historical rate series for sales taxes, which are imposed by a variety of state and local jurisdictions. In order to estimate sales tax deductions on income tax returns, we subtracted from consumption expenditures certain items which are generally exempt – rent, medical care and charitable contributions – and multiplied the remainder by 4.075%, which is the unweighted average of state sales tax rates as given in the current *State Tax Guide*. Deductions are assumed to equal zero in the four states which impose no sales tax.

Household Data. Our basic household-level data source was the 1972-73 Consumer Expenditure Survey (CES).[12] From the CES we derived annual consumption expenditure series along with information necessary to define filing status, exemptions, deductions and credits.

We began with the CES data base constructed by Robert Hagemann [1982] for his study of household-specific price indexes. We then excluded households which were located in Alaska or Hawaii, which moved or changed tenure status during the survey year, or whose primary earner was retired ~~or~~ unemployed. We also restricted our attention to individuals and households containing a husband-and-wife couple. These and other minor edits resulted in a data containing 7,351 consumer units.

Hagemann [1982] defined household-level ''market baskets'' by calculating expenditures in 37 categories of consumption. He then constructed fixed-weight price indexes by linking the base period (1972-73) expenditure weights to the appropriate CPI price series. We follow the identical procedure, differing only in adding a 38th category of consumption – charitable contributions, important because of their tax deductibility – and in extending the consumption indexes back to 1967.[13] We thus implicitly assume that all households in our sample face the same rates of inflation for individual consumption items.

For convenience, and having no CES information to the contrary, we assume that all married couples file jointly, regardless of their mix of incomes. For FICA classification purposes we defined each household as wage-and-salary or self-employed according to the classification of the chief earner. We have sufficient information on the ages of family members to determine the eligibility of households for elderly exemptions and for the Credit for the Elderly. We assume that any consumer unit containing a child met the demographic requirements for the Earned Income Credit in applicable years.

Tenure status is an important determinant of tax liability. For households who own their own homes, shelter costs were defined as the estimated market rental value of the dwelling unit. (Here we followed the Hagemann study and the treatment of housing costs which will be implemented in the CPI beginning in January 1983.) This implicit rental income is not taxed, unlike the earnings used by renters to purchase shelter services.

We also dealt with tenure status in determining itemized tax deductions. To a significant degree, deductible expenditures are not included in consumption costs. A large proportion are, instead, interest costs incurred as part of an investment or intertemporal consumption allocation decision. Wishing to abstract from such decisions here, we did not include automobile and other installment interest payments in our deduction computations. We also ignore such other non-consumption expense categories as personal property taxes, union dues and professional expenses. Deductions which are part of consumption, and which we did explicitly calculate, are health insurance, other medical care expenditures and charitable contributions. Sales taxes are incorporated in CES expenditure figures, and we, therefore, also estimated sales tax deductions as described earlier in this section.

Quantitatively, however, deductions are dominated by homeowners' mortgage interest and real property tax expenses. In order to obtain a realistic pattern of itemization in our simulations, it was therefore necessary to deal with these costs. We did not want to incorporate such investments in the income necessary to achieve a given consumption level. It would also have been incorrect to represent those households with large mortgage deductions as being, *ceteris paribus*, better off than those which owned homes outright. The solution was to compute H, the sum of mortgage interest and property taxes, and **subtract** it from the household's applicable standard deduction S. The household itemizes the remaining deductible expenses when they exceed S-H.

Households in the 1972-73 CES carried mortgages of varying ages. The price series used to index mortgage interest expenditures was a comparable seven-year unweighted average of lagged and current values of the CPI Contracted Mortgage Interest Cost series.

We note that the only CES tax or income data we require are property tax payments, for use in determining itemization behavior, and the earnings data used to categorize taxpayers as wage-and-salary or self-employed. The indexes we present in later sections are estimates of the necessary earnings and tax payments associated with observed consumption levels. They are not based on reported CES values of income, taxes withheld or refunds.

3.2 Index Computation

The fundamental computational problem in constructing our TPI series is as follows: given a value of consumption expenditures C, what is the value of gross income Y such that after-tax income equals C? The result requires knowledge of the state and federal rate structures and the categories of income to which they apply, the levels of deductible expenditures, and the details of allowable credits and standard deductions. Because the relationship between income and tax is not smooth or even monotonic, there can be no closed form solution to the problem. We shall use a simplified example to demonstrate the iterative procedure employed for this study.

Let r_F be the marginal federal income tax rate and let P_F be a constant incorporating the difference between r_F and the rates applicable in lower income brackets (i.e. P_F reflects the difference between the marginal and average tax rate). Let r_S and P_S be the corresponding state values and let r_A be the FICA contribution rate. Define C, E and D as the level of consumption, the value of exemptions, and the value of deductible consumption expenditures, respectively. Then we have

$$S = P_S + r_S(Y\text{-}E\text{-}D) \tag{1}$$

$$F = P_F + r_F(Y\text{-}E\text{-}D\text{-}S) \tag{2}$$

and

$$A = R_A Y \tag{3}$$

where S, F and A are, respectively, state tax, federal tax and FICA contributions. Note that state tax is deductible on the federal return. Since we must have

$$Y = C + S + F + A \tag{4}$$

we obtain, through straightforward algebra,

$$Y = \frac{1}{(1-r_F)(1-r_S)-r_A} [C + P_F + (1-r_F)P_S - (r_F + r_S - r_F r_S)(E + D)] \qquad (5)$$

This computed Y value is the solution if (i) Y falls in the same state and federal tax brackets used to define r_S, r_F, P_S, and P_F; (ii) Y does not lie above the ceiling FICA income value, in which case the marginal FICA rate would be zero rather than r_A; and (iii) D exceeds the value of the standard deduction at income equal to Y. If these conditions do not hold, new parameters corresponding to Y are inserted into (5) and another solution is attempted.

The procedure we followed is basically an elaboration of equation (5) to allow for tax credits, percentage and flat standard deductions, states with unusual tax formulas, the non-taxed status of implicit rental income, and other complications. It can be shown that if the piecewise linear function relating after-tax to before-tax income were concave, the above procedure would necessarily converge to a solution. Although the concavity condition does not hold, we experienced little difficulty in solving for Y. In 110,265 attempts (7,351 households for each of 15 years) the process never failed to find a solution in 10 iterations or less. We closely examined a one-in-1,000 sample of 110 cases, and in only one of these did the solution process require as many as five iterations.

The solution values of Y for each household and year provide us with the information necessary to compute our tax and price index series. We can also compute indexes of C, F, S and A to examine the separate influences of each on the inflation experience of our sample households. The results of these computations are discussed in the following sections.

4. Estimated Tax and Price Indexes

Table 1 presents our simulated annual cost series for consumption and tax, based on our sample of 7,351 CES households. The series are displayed graphically in Figure 1. The consumption cost series was derived by linking observed 1972-73 expenditure levels to CPI component price series, as described in Section 3.1 above. To compute the three tax series and the total cost series, we inserted each household's consumption level into the algorithm of Section 3.2, using the tax parameters for the appropriate year, state and filing status.

In computing sample means, household expenditures were weighted by their CES sampling weights. This follows the official CPI methodology for estimating population expenditure totals.

The figures in the table highlight the expanding share of taxes in our simulated household budgets. In 1967, the sample consumer units required an average of $8,930 in gross income in order to retain after-tax funds sufficient to purchase their 1972-73 observed consumption bundles. Of this $8,930, approximately $813, or 9.1%, was allocated to federal income tax. Another 0.9% and 2.8% went for state income tax and Social Security contributions, respectively, leaving slightly more than 87% for consumption goods.

By 1981, the mean direct consumption cost of $19,498 was only 78% of the income required to achieve that level of consumption expenditure. Federal, state and FICA taxation all rose more rapidly than the prices of market goods. These three tax components totalled, respectively, 14.0, 2.7 and 5.3% of required gross income.

Table 2 presents index series derived from the value series in Table 1, with the indexes set at 100 in the base year of 1967. As expected given its method of calculation, our estimated CPI closely approximates the published CPI-U,X1 series with the same rental equivalence definition of homeowner shelter cost. Again, we see that all three tax series rose much more rapidly than goods and services prices over the study period. In 1981 the TPI index stood at 280.1, 29.7 index points above the consumption cost index. This indicates that had the before-tax earnings of sample households been indexed by the CPI, these households on average would have fallen far short of retaining the same purchasing power in 1981 that they had in 1967.

Year-to-year percentage changes in the TPI and component indexes are shown in Table 3. The table shows again that the TPI rose by an average 0.8% per year faster than the CPI. The widest differences were in the last two years of the study period. Between 1979 and 1981 the cost of our fixed market basket rose by 22.1%, while the income necessary to purchase it rose by 27.5%.

The years of greatest inflation in goods prices were 1974, 1979, 1980 and 1981. These

are also the four years in which our TPI increases at a double-digit rate. The component tax indexes, too, tend to increase with inflation, especially given the ''bracket creep'' arising from progressive state and federal income tax schedules. However, each tax series also reflects changes in marginal rates and other system parameters. For example, the imposition of the tax surcharge for nine months in 1968 and the full year 1969 contributed to sharp increases in the federal tax index in those years. As noted in Section 3, the surcharge was phased out in 1970 and a lower rate schedule for single taxpayers was introduced in 1971. These and other IRS changes resulted in three consecutive years of declining federal tax liabilities. The introduction of the Personal Exemption Credit and Earned Income Credit in 1975, and the Credit for the Elderly and Taxable Income Credit in 1976, helped make those years of relatively low index change. Finally, our federal tax index rises more slowly in 1979 than in 1978, despite the accelerating rate of price inflation, because of the new lower marginal rate schedules.

The state tax index rises at the greatest annual rate, 16.4% over our 1967-81 year study period. It should be noted that this rapid growth relative to the other tax and price series tends to be concentrated in the early years, when several states were imposing income taxes for the first time. As shown in Table 2, the state tax liability rose 110.7% between 1967 and 1972, while consumption costs and federal taxes each rose only 21.5%. By contrast, in five of the nine years after 1972 the federal tax index rises more rapidly than the state index.

Only one sample year, 1970, saw no change in either the FICA tax rates or the earnings ceiling. Correspondingly, our FICA contributions index rises only 1.6% in 1970, as compared to its 14-year average annual increase of 12.7%. The years of sharpest FICA index change are 1973, when the wage-and-salary contributions rate was raised by 12.5%, the self-employed rate by 6.7% and the ceiling by 20% and 1981, when the rate increases were 8.5% and 14.8%, respectively, and the ceiling rose by 14.7%.

In Table 4 we classify sample households into consumption quartiles. A household's ''real'' consumption is defined as the simple average of the 1967, 1974 and 1981 values of its base year (1972 or 1973) consumption bundle. In defining quartiles of real consumption, households are again weighted by CES sample weights. The table displays, for each

quartile, the 1967 and 1981 mean consumption costs and tax payments, along with the 1981 estimated tax and price index levels.

Two primary conclusions can be drawn from Table 4. First, the consumption cost indexes are similar across quartiles. That is, it does not appear that inflation rates have been markedly greater or lower for the types of goods and services purchased by high-consumption households. The second conclusion is that when we include taxes in total cost, we observe a positive correlation between inflation and real consumption.

Tax inflation has had a greater effect on high-consumption households, both because the weight of taxes in their budget is larger and because their taxes have increased more rapidly. For example, as shown in Table 4, federal income tax, state income tax and FICA contributions had shares of 11.1, 1.2 and 2.0% in the 1967 costs of households in quartile IV, the highest consumption quartile. By contrast, in 1967 federal and state income tax combined for only a 7.2% share of quartile I budgets. (Since the FICA schedule is characterized by a flat rate and a ceiling, the FICA share declines with income.) Partially as a result of the several new low-income credit programs, the annual rate of federal tax inflation was only 7.7% in quartile I, compared to 10.1, 10.7 and 11.8% for quartiles II, III and IV. A similar relationship holds for FICA contributions. This follows from the several increases in the contributions ceiling, which primarily affect high-income earners. Although the state tax index rises more slowly in the higher-consumption quartiles, state taxes continue to consume a much greater share of total cost for those households.

In Table 5 we present consumption and tax expenditure levels and index values in the manner of Table 4, broken down by several household characteristics. Once again, 1967-81 price inflation rates were relatively similar across groups, while the total cost index is higher for joint filers and self-employed workers, and slightly higher for homeowners and younger households. To some extent these differences reflect relative income effects. Married couples, for example, tend to have higher consumption levels than single individuals and so are more seriously affected by increases in the tax indexes. In other cases, changes in tax laws have been significant. The introduction of the Credit for the Elderly caused the federal tax index for households with heads over 64 years old to drop by 10.8% in 1976, while the corresponding index for younger households rose 7.7%. The federal tax index fell 12.0%

for single filers in 1971, when the lower rate schedule was imposed, compared to only a 1.4% drop for married couples filing jointly.

5. Analysis of the Variation in Tax Liabilities as a Percentage of Income

The calculation of individual household TPI's provides a rich data base for analyzing the relationship between tax liabilities and household characteristics at the micro level. In this section we will focus on a limited number of household characteristics on which we regress both component and total tax liabilities as a percentage of total income. Our primary interest is in the relationship between consumption levels and tax percentages, but we also include a small number of variables in each regression which have an explicit role in the tax regulations. In so doing, we enhance the analysis of the previous section by analyzing the relationship between taxes and a particular household characteristic conditional on the levels of the remaining characteristics. Regressions were run for both 1967 and 1981 for each filing status. The years were chosen to facilitate analysis of changes in both percentage liabilities and schedule progressivity over our sample period. We partitioned by filing status not only to account for differences in rate schedules but also because this allows us to have comparability in consumption per adult within each regression. For the state tax regressions we exclude households in states without a state tax in the appropriate year.

The regression results are presented in Tables 6 and 7. With the exception of consumption, the regressors are self-explanatory. Consumption is represented by the logarithm of each household's fixed consumption set evaluated at the average of 1967, 1974 and 1981 prices. (For ease in interpreting coefficients, the logarithm of consumption level was standardized to have a mean of zero and standard deviation of one within each filing status subgroup.) We emphasize the impact of (log) consumption rather than income level on tax burden since (1) it is consumption which determines satisfaction level and (2) consumption is exogenous in our analysis. The dependent variable in each regression is the specified tax component measured as a percentage of total income.

The intercept in each equation is the predicted percentage of total income accounted for by each tax component for a "reference" household with subgroup mean consumption and a value of zero for each of the dummy variables included in the equation. Looking at changes in the intercepts between 1967 and 1981, it is quite apparent that the tax

percentages for state and FICA taxes have increased drastically. The federal tax percentage increased at a more moderate rate for reference households filing a joint return and by only 1.1 percentage points over the entire period for reference, single-person households. As expected, elderly and dependent members of the household have a negative impact on tax percentages, though this impact did not change drastically between periods. Also as expected, the direct impact of self-employment status on FICA taxes and the indirect impact on overall taxes increased substantially over the sample period, as the differential between the self-employed and wage and salary contribution rates widened. The significantly negative coefficients in the 1981 state and total tax regressions on the dummy variable denoting that the household is in a state which did not have a state tax in 1967 indicate that states with ''new'' state taxes have not caught up to states with more established state taxes in terms of the percentage of income accounted for by the state tax component. The coefficient on renter status reflects the fact that owner-occupants have substantial, non-taxed income in kind.

Of equal interest to changes in the level of taxes are changes in the progressivity of the tax system. Before describing these changes, however, it is important to emphasize that we calculate necessary income and tax burdens without fully modelling the impact of tax rules on the financial behavior of households. Consequently, our estimates of progressivity are hypothetical and refer to the progressivity which would only apply if households did not tailor their income, investment and consumption decisions to the tax system. Actual progressivity, which includes the impact of this behavior, could be substantially less.

In Figures 2 through 5 we plot the relationship between the standardized logarithm of consumption and percentage tax liability for each filing status for each of the tax components as well as total taxes. Of most interest is the change in the relationship between federal taxes and consumption (Figure 2). For both filing statuses the progressivity of federal taxes increased substantially. Even though the average percentage liability increased, for very low consumption level families and for most single persons with consumption levels below the geometric mean the percentage liability went down between 1967 and 1981. State taxes, displayed in Figure 3, show little change in progressivity for either filing status, the primary change being an upward shift. The change in the relationship between the social

security contributions percentage and consumption, displayed in Figure 4, differs substantially by filing status. For single-person households the relationship became flatter while for families the curvature increased. Finally, Figure 5 displays the relationship between total tax liabilities and consumption level. Because of the importance of federal tax in the overall tax structure, these relationships are dominated by the relationship between it and consumption.

6. Summary and Directions for Future Research

The purpose of this paper has been to define and estimate a "tax and price" index which incorporates direct as well as indirect taxes. Current U.S. CPI methodology measures changes in the minimum expenditure necessary to consume a fixed set of goods and services and consequently approximates an expenditure-based cost-of-living index. The indexes we compute measure changes in the minimum income, before taxes, necessary to yield the purchasing power required for the same fixed set of goods and services and approximate, in an analogous fashion, an income-based cost-of-living index.

Our tax and price indexes were calculated at the individual household level and used detailed procedures to add federal, state and Social Security components to an expenditure-based "CPI". All of the tax components increased substantially faster over the sample period than the estimated CPI, with state taxes increasing at the fastest rate. Partitioning the sample by consumption quartile, we found that, while the CPI component varies little among quartiles, the rate of increase in taxes has been positively related to consumption level.

We also analyzed the relationship between tax liabilities and a limited number of household characteristics within a multiple regression framework. This analysis indicates that, *ceteris paribus*, tax liabilities as a percentage of income have increased, though the federal tax percentage increased relatively little, especially for single-person households. The "progressivity", with respect to consumption, of the federal system also increased substantially over the sample period.

Our future plans include extensions of this research in two directions. First, additional data will be acquired in order to improve the accuracy of our tax and price index and to

extend its coverage to other years, income categories and household types. Examples of these data are local area income tax schedules, state sales tax rates, federal schedules for heads of households and married couples filing separately, and IRS regulations regarding retirement income, interest, dividends and capital gains.

Concurrently with the incorporation of new data, we plan to begin employment of our indexes in analytical work. We will use the TPI as a means of examining such phenomena as ''bracket creep'', the Kennedy and Reagan tax cuts, and the indexation of wage rates and entitlement programs. Since we have computed our indexes at the household level, we can determine the extent to which the tax structure may have led to changes in the relative economic well-being of particular population subgroups, defined by, for example, income source or tenure status. Finally, the tax rate data and computational programs employed here will also be used to construct important related economic series, such as the rate of return to investments in owner-occupied housing.

TABLE 1. Annual Consumption Costs and Tax Liabilities, 1967 to 1981

Year	Sample Mean Values				
	Cost of Consumption	Federal Tax	State Tax	FICA Contribution	Total Cost
1967	7,786	813	81	250	8,930
1968	8,069	960	96	281	9,406
1969	8,408	1,076	117	315	9,916
1970	8,799	1,037	126	320	10,282
1971	9,182	1,011	148	351	10,692
1972	9,460	988	170	386	11,004
1973	10,064	1,144	196	497	11,901
1974	11,120	1,418	242	584	13,364
1975	12,022	1,464	282	625	14,394
1976	12,722	1,565	320	668	15,275
1977	13,539	1,695	361	713	16,309
1978	14,488	1,989	414	797	17,688
1979	15,967	2,231	470	952	19,621
1980	17,801	2,867	570	1,085	22,323
1981	19,498	3,506	675	1,334	25,013

TABLE 2. Tax and Price Indexes, 1967 to 1981

Year	TPI Component				
	CPI	Federal Tax	State Tax	FICA Contribution	TPI
1967	100.0	100.0	100.0	100.0	100.0
1968	103.6	118.1	119.2	112.4	105.3
1969	108.0	132.3	145.3	125.7	111.0
1970	113.0	127.6	156.2	127.8	115.1
1971	117.9	124.4	183.3	140.4	119.7
1972	121.5	121.5	210.7	154.4	123.2
1973	129.2	140.7	243.8	198.7	133.3
1974	142.8	174.5	300.0	233.3	149.7
1975	154.4	180.2	350.3	249.8	161.2
1976	163.4	192.6	397.6	266.7	171.1
1977	173.9	208.6	448.6	285.0	182.6
1978	186.1	244.7	514.0	318.5	198.1
1979	205.1	274.5	583.9	380.4	219.7
1980	228.6	352.7	707.4	433.4	250.0
1981	250.4	431.3	838.3	533.0	280.1

TABLE 3. Percentage Index Changes, 1967 to 1981

Year	TPI Component				
	CPI	Federal Tax	State Tax	FICA Contribution	TPI
1967-68	3.6	18.1	19.2	12.4	5.3
1968-69	4.2	12.0	21.9	11.8	5.4
1969-70	4.7	−3.6	7.5	1.6	3.7
1970-71	4.3	−2.5	17.4	9.9	4.0
1971-72	3.0	−2.3	15.0	10.0	2.9
1972-73	6.4	15.8	15.7	28.7	8.2
1973-74	10.5	24.0	23.0	17.5	12.3
1974-75	8.1	3.2	16.8	7.1	7.7
1975-76	5.8	6.9	13.5	6.8	6.1
1976-77	6.4	8.3	12.8	6.9	6.8
1977-78	7.0	17.3	14.6	11.8	8.5
1978-79	10.2	12.1	13.6	19.4	10.9
1979-80	11.5	28.5	21.1	13.9	13.8
1980-81	9.5	22.3	18.5	23.0	12.1
Average Annual Rate 1967-81	6.8	11.0	16.4	12.7	7.6

TABLE 4. Annual Consumption Costs and Tax Liabilities By Consumption Quartile

Quartile	Cost of Consumption	Federal Tax	State Tax	FICA Contribution	Total Cost
I 1967	3,655	277	20	161	4,112
1981	9,287	780	189	665	10,921
Index	251.4	281.8	949.0	412.9	265.6
II 1967	6,107	495	40	249	6,891
1981	15,479	1,900	383	1,120	18,881
Index	253.5	383.9	953.4	449.5	274.0
III 1967	8,347	799	72	292	9,508
1981	20,941	3,313	648	1,566	26,468
Index	250.9	414.9	906.5	537.0	278.4
IV 1967	13,034	1,681	190	300	15,205
1981	32,282	8,027	1,480	1,987	43,775
Index	247.7	477.6	778.9	663.0	287.9

TABLE 5. **Annual Consumption Costs and Tax Liabilities By Household Characteristics**

Household Characteristic		Cost of Consumption	Federal Tax	State Tax	FICA Contribution	Total Cost
Filing Status						
Single	1967	4,237	634	62	182	5,115
	1981	10,404	2,030	410	819	13,663
	Index	245.6	320.1	659.8	450.4	267.1
Joint	1967	8,339	841	83	261	9,524
	1981	20,914	3,735	716	1,414	26,781
	Index	250.8	444.4	859.0	542.0	281.2
Income Source						
Wage and Salary	1967	7,697	797	79	244	8,818
	1981	19,282	3,418	665	1,297	24,661
	Index	250.5	428.7	841.7	531.5	279.7
Self-Employed	1967	9,536	1,117	111	375	11,140
	1981	23,758	5,223	879	2,075	31,935
	Index	249.1	467.6	791.3	552.8	286.7
Tenure						
Owner	1967	8,565	821	76	256	9,718
	1981	21,478	3,744	680	1,399	27,300
	Index	250.8	455.8	893.2	547.6	280.9
Renter	1967	5,749	790	92	237	6,868
	1981	14,321	2,882	664	1,165	19,032
	Index	249.1	364.6	719.9	491.8	277.1
Age of Head						
Under 65	1967	7,963	846	84	255	9,148
	1981	19,930	3,647	696	1,366	25,640
	Index	250.3	431.2	829.5	535.5	280.3
65 or Over	1967	5,702	424	40	194	6,361
	1981	14,406	1,838	426	956	17,627
	Index	252.6	433.2	1,054.1	493.5	277.1

TABLE 6. Estimated Parameters for Percentage Tax Liability Regressions, Joint Filing Status (Standard errors in parentheses)

Variable	Federal Tax		State Tax*		FICA Contributions		Total Taxes	
	1967	1981	1967	1981	1967	1981	1967	1981
Intercept	9.080 (.037)	12.490 (.051)	1.261 (.032)	3.066 (.037)	3.069 (.006)	5.592 (.009)	13.189 (.053)	20.861 (.074)
Log(Consumption)	3.230 (.023)	5.596 (.031)	.774 (.019)	.971 (.021)	−.666 (.005)	−.484 (.007)	3.004 (.029)	5.908 (.041)
[Log (Consumption)]^2	.112 (.014)	.215 (.019)	.065 (.012)	.045 (.013)	−.156 (.003)	−.279 (.004)	.006 (.017)	−.007 (.025)
Number of Dependents	−1.142 (.013)	−.990 (.017)	−.140 (.011)	−.111 (.012)			−1.188 (.016)	−1.036 (.022)
Number of Elderly	−1.591 (.050)	−1.462 (.068)	−.236 (.042)	−.158 (.046)			−1.771 (.062)	−1.618 (.088)
Self-employed					1.254 (.020)	2.030 (.029)	.823 (.118)	1.667 (.168)
No State Tax in 1967				−1.388 (.041)			−.833 (.053)	−.959 (.086)
No State Tax in 1981								−1.601 (.116)
Renter	3.420 (.054)	3.290 (.073)	.897 (.045)	1.176 (.049)	.498 (.011)	.901 (.017)	4.440 (.067)	5.208 (.095)
R-square	.799	.842	.337	.395	.832	.729	.728	.790
Sample Size	6,335	6,335	3,519	5,313	6,335	6,335	6,335	6,335

TABLE 7. Estimated Parameters for Percentage Tax Liability Regressions, Single Person Filing Status (Standard errors in parentheses)

Variable	Federal Tax		State Tax*		FICA Contributions		Total Taxes	
	1967	1981	1967	1981	1967	1981	1967	1981
Intercept	7.573	8.263	.831	2.197	3.235	5.166	11.987	15.786
	(.091)	(.021)	(.084)	(.093)	(.025)	(.029)	(.136)	(.186)
Log(Consumption)	3.289	5.631	.773	1.028	−.293	−.079	3.444	6.412
	(.049)	(.065)	(.044)	(.048)	(.014)	(.016)	(.062)	(.091)
[Log (Consumption)]^2	−.040	.402	.077	.127	−.167	−.092	−.147	.431
	(.032)	(.043)	(.028)	(.032)	(.009)	(.011)	(.044)	(.061)
Elderly	−3.995	−3.823	−.496	−.340			−4.371	−4.148
	(.135)	(.180)	(.123)	(.133)			(.185)	(.253)
Self-employed					1.360	2.162	1.745	3.224
					(.080)	(.093)	(.376)	(.515)
No State Tax in 1967				−1.365			−1.635	−1.391
				(.098)			(.129)	(.205)
No State Tax in 1981								−2.085
								(.276)
Renter	5.367	4.787	1.145	1.407	1.076	1.508	7.085	7.499
	(.101)	(.134)	(.093)	(.098)	(.029)	(.033)	(.138)	(.189)
R-square	.898	.907	.452	.503	.714	.722	.863	.878
Sample Size	1,016	1,016	579	852	1,016	1,016	1,016	1,016

*Sample limited to observations in states with state income tax.

Footnotes

[1] Cost-of-living subindexes are described in Pollak [1975]. The relationship between the CPI anc a cost-of-living subindex is outlined in Gillingham [1974].

[2] The motivation and methodology for the United Kingdom's Tax and Price Index are describec in U.K. Central Statistical Office [1979].

[3] The U.S. Bureau of Labor Statistics currently produces a number of alternative CPIs, including an experimental measure using a flow-of-services approach to measuring shelter costs. We adopt this approach, described in more detail in Gillingham [1983], in this paper.

[4] Under restrictive separability conditions, the CPI can also be interpreted as an approximatior to a partial subindex which is independent of the levels of all of the variables assumed fixed ir the conditional subindex. See Pollak [1975] and Gillingham [1974].

[5] Presumably, it would be easy to treat savings as another good the "price" of which is representec by the household's CPI. We have not followed this approach because it would be unlikely tc change the qualitative results and because estimates of savings available from the CES are extremely unreliable.

[6] This is certainly an inadequate solution for this problem for both a CPI and TPI. However, a: an interim fix, the problems are mitigated by the fact that the importance of non-housing durable in the CPI is small relative to shelter.

[7] Housing investment expenses are an important determinant of whether a household itemizes it: deductions. To obtain a representative proportion of "itemizers" we adopt additional conventions in the next section for determining whether a household itemizes.

[8] It should be noted that this general description of the federal tax system masks several change in the terminology used on IRS forms. For example, the standard deduction has at times beer represented only implicitly, through tax tables, and has sometimes been referred to as the "lov income allowance" or the "zero bracket amount".

[9] In this paper we deal only with single taxpayers and with married couples filing jointly. Separat rate schedules and deduction amounts have applied to married persons filing separately or to taxpayers filing as heads of households.

[10] We ignore direct taxes applied to interest income, dividends, and other categories of "non-earned" income.

[11] The only exception is the state of Maryland, where we take into account that every county represented in our household sample imposes a 50% surcharge on the state liability.

[12] The 1972-73 CES survey is described by Carlson [1974].

[13] CPI price series are available for years prior to 1967 only at a more aggregated level. We usec the all-items CPI (specifically, the X1 series employing the rental equivalence measure o homeownership cost) as a price series for charitable contributions, on the basis that household are contributing units of generalized purchasing power.

* The authors wish to thank Ernst Berndt, Robert Hagemann, and participants in the Conference fo helpful comments and suggestions. They would also like to express their particular thanks to Rober Hagemann for graciously making available his data and computer programs.

References

Carlson, Michael D., "The 1972-73 Consumer Expenditure Survey", *Monthly Labor Review*, 97 (December 1974), pp.16-23.

Commerce Clearing House, *State Tax Guide*, 1982.

_________ *State Tax Handbook*, various issues.

Gillingham, Robert, "A Conceptual Framework for the Consumer Price Index", *Proceedings*, American Statistical Association Business and Economic Statistics Section, 1974.

_________ "Measuring the Cost of Shelter for Homeowners: Theoretical and Empirical Considerations", *Review of Economics and Statistics,* 65 (May 1983), pp. 254-265.

Hagemann, Robert P., "The Variability of Inflation Rates Across Household Types", *Journal of Money, Credit, and Banking*, 14 (November 1982), pp.494-510.

Pollak, Robert, "Treatment of Taxes in the Consumer Price Index ", BLS Mimeo, 1972.

_________ "Subindexes of the Cost of Living Index", *International Economic Review*, 16 (February 1975), pp.135-150.

U.K. Central Statistical Office, "The tax and price index – sources and methods", *Economic Trends*, No. 310 (August 1979), pp.81-89.

U.S. Department of the Treasury, *Individual Income Tax Returns*, Internal Revenue Service Publication 79, various updates.

_________ *Your Federal Income Tax*, Internal Revenue Service Publication 17, 1980 and 1981 updates.

Figure 1

Tax and Price Component Series 1967-1981

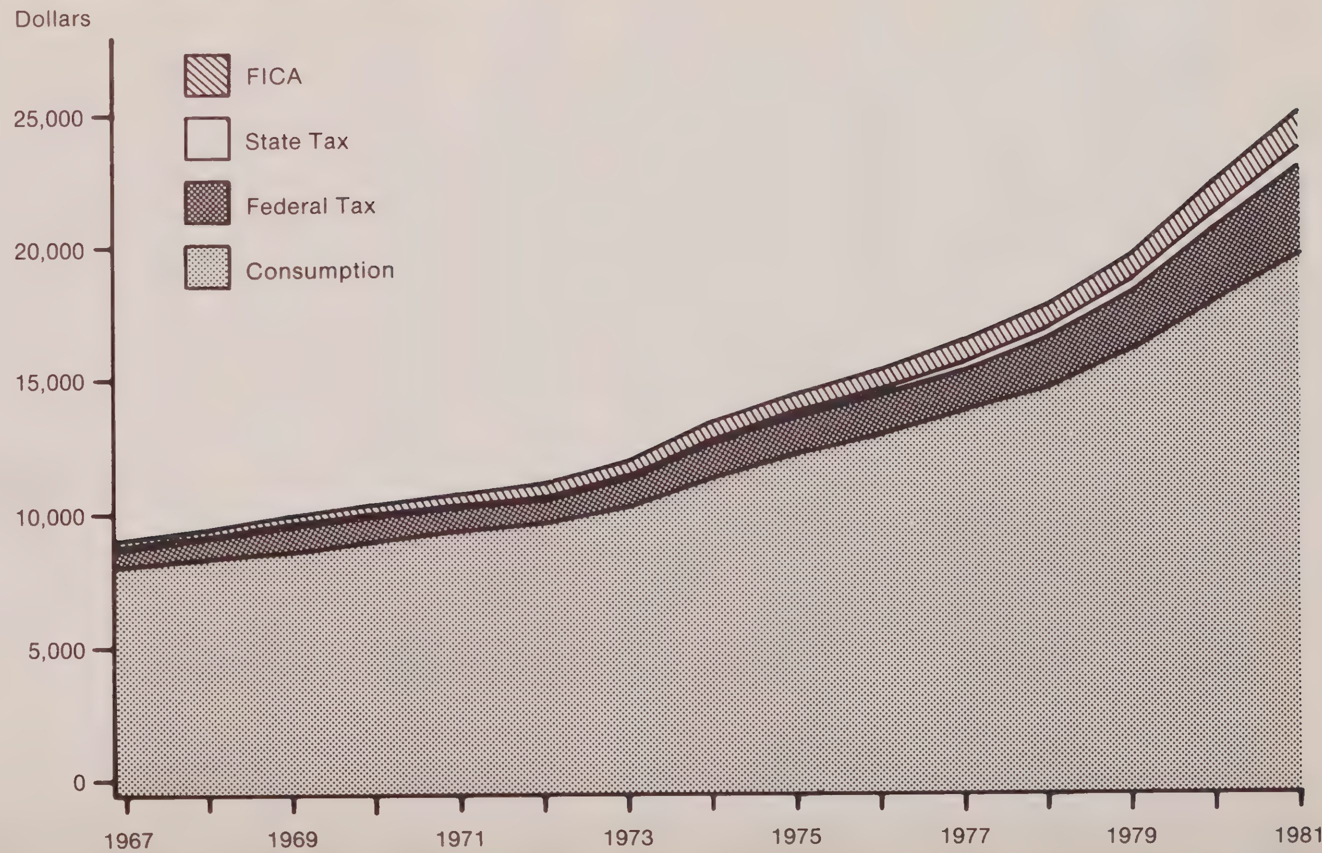


Figure 2

Federal Tax as a Percentage of Income 1967 and 1981

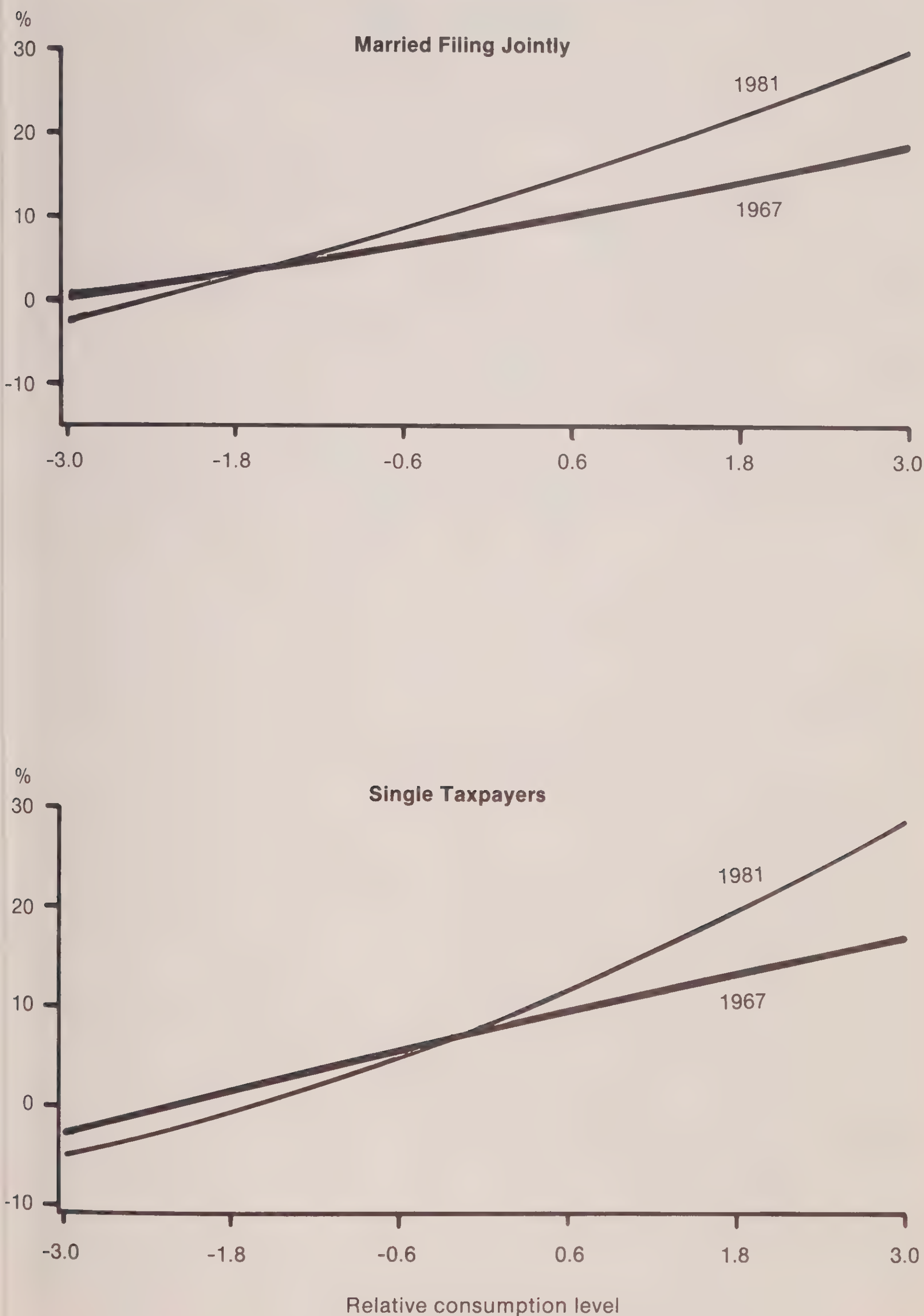


Figure 3

State Tax as a Percentage of Income 1967 and 1981

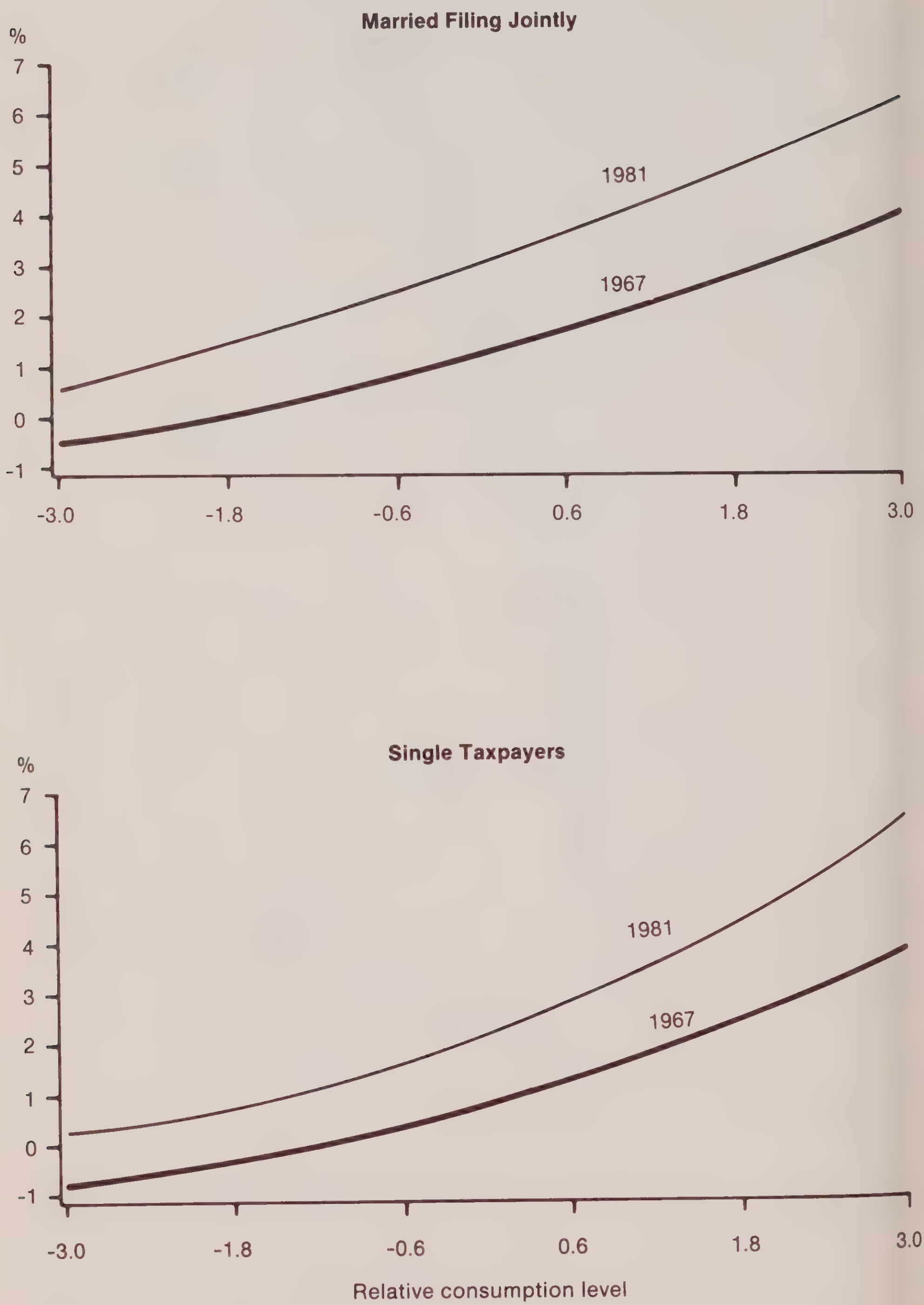


Figure 4

FICA Payments as a Percentage of Income 1967 and 1981

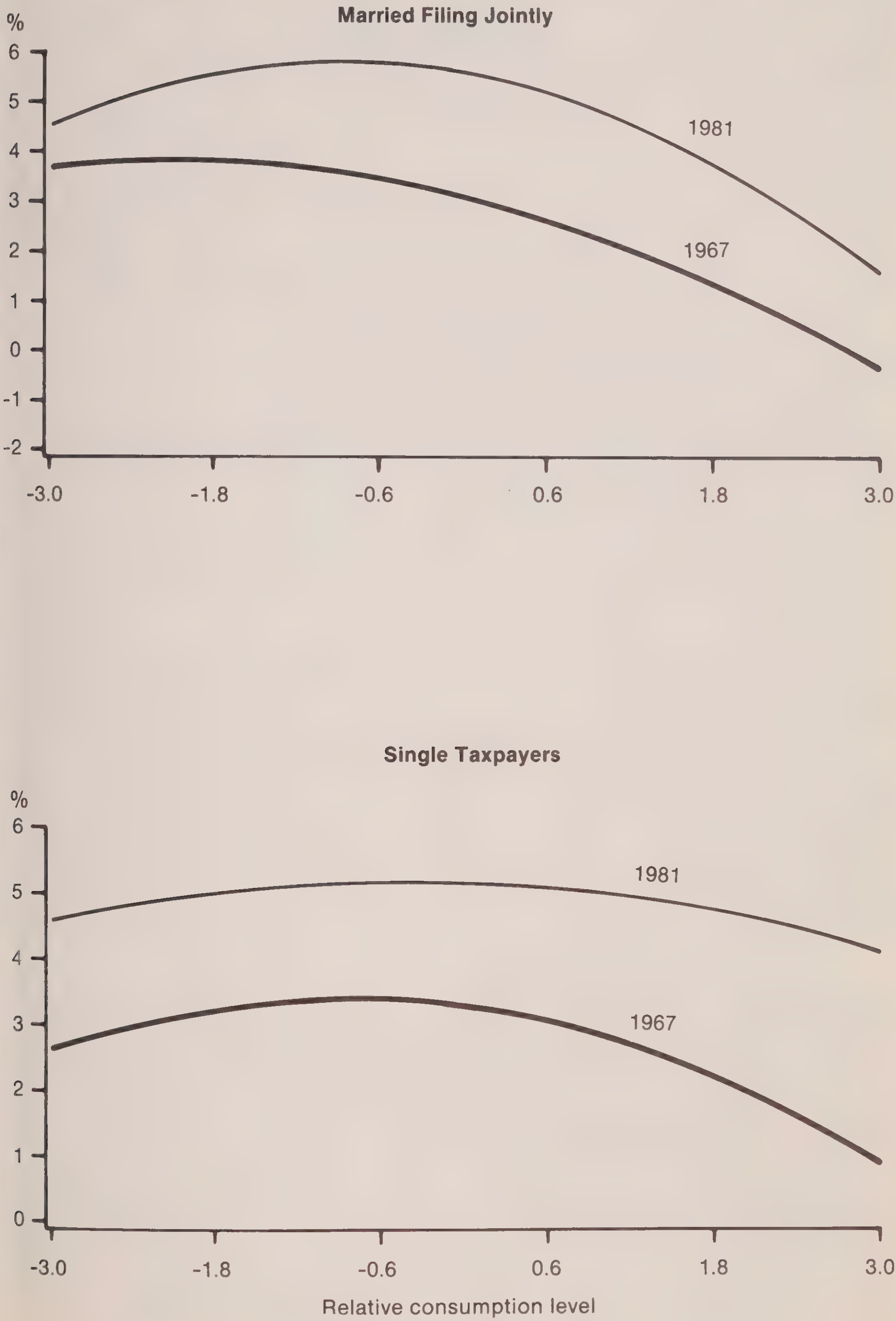
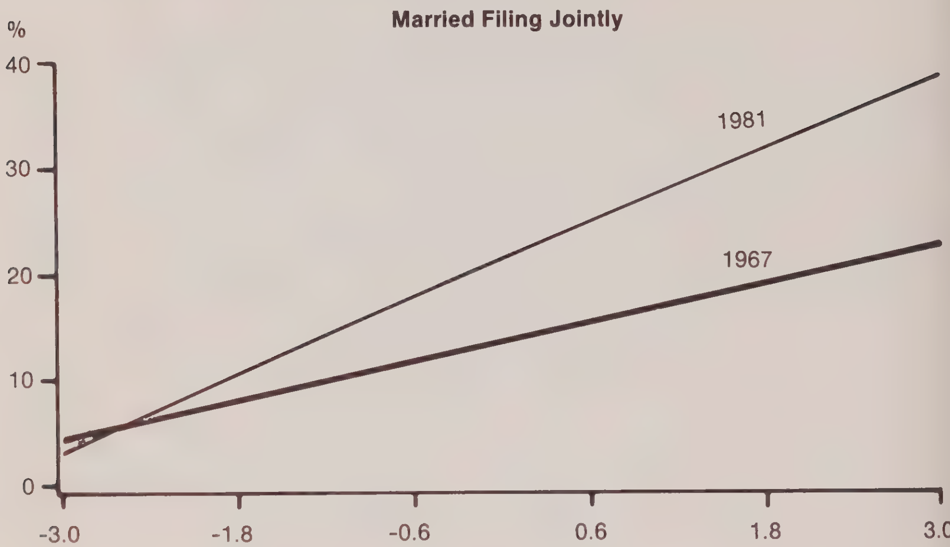
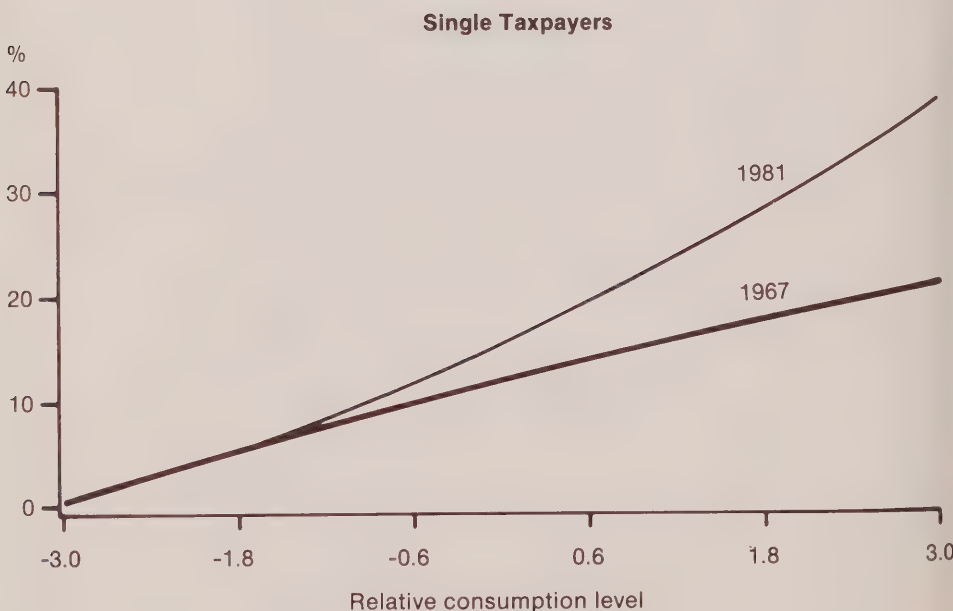


Figure 5

Total Tax as a Percentage of Income 1967 and 1981





PUBLIC GOODS AND PRICE INDEXES

Pour vous fournir une version dans la langue officielle de votre choix, le texte anglais est suivi du texte français (p.679) dans cette publication.

Claude Montmarquette
Université de Montréal

SUMMARY

The purpose of this paper is to incorporate public goods into the calculation of a constant utility price index covering all commodity prices in the economy. The study links public goods to the traditional literature on price measurement using the median voter model, borrowed from public choice literature. Various simulations and empirical results of the proposed index are presented using Canadian data for the period of 1948 to 1980.

As noted in a reference paper published by Statistics Canada [1982, p.9], the Consumer Price Index measures ''pure'' price movement only. More specifically, the CPI measures price changes by comparing, through time, the cost of a basket of goods of fixed or equivalent quantity and quality specified according to purchases made by the target population. Statistics Canada stresses the ''pure'' aspect of price movement because, over and above the very real difficulties involved in calculating the CPI, the use of aggregate population data biases the results because it fails to present users with valid alternatives.

In the case of the CPI, the use of this index as a cost-of-living measure for the indexation of rents, assets or other goods and services, wages and other income has received a great deal of attention from researchers in studies aimed at defining the limits, biases or conditions associated with the use of the CPI for such purposes. Economists have often challenged this measure of price movement because of its arbitrary confinement to a specific group of individuals and a given group of goods and services. The idea of updating the index based on more recent expenditures to reflect changes in the quality of the goods and services, and spatial indexes and sub-index theory, to mention but a few areas of interest,

represents a partial response to this criticism. However, one continuing major shortcoming of the literature is that the role of public goods in the production of a price index has been largely ignored. The literature has remained virtually mute on this question since the discussion of Kessel [1961] on the role of direct and indirect taxes in the calculation of the CPI.[1] Government is currently responsible for some 40% of total expenditure in the economy, yet the CPI as currently calculated makes allowance for only a very small portion of this expenditure. To be at all relevant, a price index must make sufficient allowance for public goods.

The purpose of this paper is to incorporate public goods into the calculation of a constant utility price index covering all commodity prices in the economy.

The first section of the study will link public goods to the traditional literature on price measurement using the median voter model, borrowed from public choice literature.

The second section will define a constant utility price index including the goods produced by governments. On an abstract level, certain dimensions of this index are described, attesting to the value and relevance of this index. This index is also compared with the indexes currently in use. Finally the problems associated with direct and indirect taxes in defining the index are discussed.

The third section presents various simulations and empirical results on this index using Canadian data for the period of 1948 to 1980.

The fourth and final section underlines the theoretical relevance of the proposed index and its operational feasibility. More generally, the approach selected in this study breaks new ground in the examination of questions on price indexes and public sector output.

1. The Median Voter Model

The incorporation of government goods in the calculation of a consumer price index poses the problem of distinguishing between quantity and price. Government services and goods are generally not priced, and the concept of public goods as goods for common consumption is not likely to facilitate their measurement. In fact, linking the price index

with a true cost-of-living measure necessitates the inclusion of all goods in an economy in a preference function. In a public goods context, the traditional problem of free riders complicates the construction of a preference function for this analysis.

The median voter model found in the literature on public choice theory (see, for example, Bergstrom and Goodman [1973], Borcherding and Deacon [1972], Dudley and Montmarquette [1981]), offers an interesting solution to these difficulties. Under certain conditions, the median voter model as initially formulated by Hotelling [1929] and Bowen [1943] shows that it is median voter preferences which govern public choices. It follows, as very clearly explained by Borcherding and Deacon [1972, p.892], that *ex post*, we are able to observe the *ex ante* preferences of the median voter, freeing us from the problem of observing the preferences of a free rider.

The median voter model has not gained universal acceptance in public choice theory; at the present time opinion is divided as to its value and relevance.[2] It is nevertheless reasonable to assume that this model is capable of providing a valid representation of reality. Moreover, for use in price index theory, the median voter model has the advantage of treating all goods equally.[3]

Let us consider the following utility median voter function:

$$U\,(X,H) \tag{1}$$

where X is a composite private commodity and H is a commodity produced by the government.

Let:

$$H = \frac{G}{qN^{\alpha}} \tag{2}$$

where G represents total government expenditure. The unit of measurement of H is defined in such a way that its unit cost is equal to q. N represents the population and α a crowding

parameter. If $\alpha = 0$, then H is a pure public good: the median voter and each individual in the economy consume the total output H. If α equals unity ($\alpha = 1$), H is the equivalent of a private good so long as each individual in the economy consumes only 1/Nth part of the total output of H.

The median voter budget constraint is defined as:

$$pX + tpX + T = m \tag{3}$$

where p is the unit price of private goods X and t is the indirect taxation rate in force, so that tpX is the value of the indirect taxes paid by the median voter; T is the value of the direct taxes paid by the median voter; and m is the income of the median voter.

The following establishes the existence of a relationship between the direct taxes paid by the median voter and government expenditure:

$$T = \tau(\,)\,G \tag{4}$$

where τ represents the proportion of total direct taxes paid by the median voter. This proportion **is a function** of government taxation structure, of the amount of government expenditure financed by indirect taxes and of budget deficits or surpluses.[4]

Substitute (2) and (4) in (3) and maximize the Lagrangian formed from utility function (1) and the redefined budget constraint. In addition to this constraint, this yields the usua equilibrium conditions which, expressed as a ratio of marginal utilities U_x and U_h, give the following equation:[5]

$$\frac{U_X}{U_H} = \frac{p\,(1 + t)}{\tau q N^{\alpha}} \tag{5}$$

At the equilibrium point, the marginal utilities ratio equals the prices ratio; it is therefore clear that the expression $\tau q N^{\alpha}$ identifies the price of H for the median voter. It will be noted that this price depends on the proportion of total direct taxes paid by the median

voter, whereas indirect taxes affect only the price of private goods; in other words, all things being equal, an indirect tax increase will (paradoxically) reduce the relative price of public goods and hence, theoretically, increase the demand for them. This point seems to have eluded the analysts of public choice theory. To paraphrase Bergstrom and Goodman [1973, p.282], if α is close to 0, there will be substantial savings for an economy with a large population, since there will be substantial savings for an economy with a large population, since there will be more consumers to share the costs of public goods without any one suffering from crowding. Moreover, if α remains close to unity, the profits earned from cost sharing between consumers will be counterbalanced by their need to share government goods.

This model can thus be used in conjunction with various standard procedures such as the integration of demand functions of the Klein and Rubin [1948-1949] type, Goldberger's [1967] indirect utility function approach or an expenditure function of the type used by Lloyd [1975] to define a true cost-of-living index. This index corresponds to the ratio of the consumer's income that leaves him on the same indifference curve regardless of the price changes taking place between the two periods, to his income during the reference period.

2. Price Index Incorporating Public Property

Let us specify utility function (1) as a CES function:

$$U = [(1-\delta)H^{\eta} + \delta X^{\eta}]^{1/\eta} \tag{6}$$

where δ is a parameter and the elasticity of substitution between private and public goods is defined as $\sigma = \frac{1}{1\text{-}n}$. Selecting the median voter budget constraint and using this CES utility function which includes public goods, an application of the Lloyd [1975] expenditure functions defines a constant utility price index between two periods:

$$T^{I} = 100\left[\frac{p^{0}(1+t^{0})X^{0}}{m^{0}}\left(\frac{p^{1}(1+t^{1})}{p^{0}(1+t^{0})}\right)^{1-\sigma} + \frac{T^{0}}{m^{0}}\left(\frac{\tau^{1}q^{1}N^{1\ \alpha^{1}}}{\tau^{0}q^{0}N^{0\ \alpha^{0}}}\right)^{1-\sigma}\right]^{\frac{1}{1-\sigma}} \tag{7}$$

where the indexes ^{0} and ^{1} identify the base period and another period respectively; T^0 corresponds to the amount of direct taxes paid by the median voter, since the price of the public goods multiplied by the quality requested yields these taxes: $\tau q N^{\alpha} \cdot H = \tau G$ by equation(2), which ultimately equals T by equation (4); m^0 represents the median voter's nominal gross income for the reference period and α is the partial and constant elasticity of substitution between private and public goods.

As Lloyd [1975] showed in the case of a homothetic function such as the CES, the constant utility price index is independent of the initial income level, since the individual commodity proportion of total expenditure essentially depends on the relative price level.[6] The index is a weighted average of order $(1-\sigma)$ in commodity prices, with the weights representing the respective proportions spent on each commodity in the reference period.

One of the advantages of this index, also noted by Lloyd, is the specific case where $\alpha = 0$: the index becomes a weighted arithmetic average, i.e. a Laspeyres index.

$$T^{I}_{Laspeyres} = 100 \left[\frac{p^1 (1+t^1) X^0}{m^0} + \frac{T^0}{m^0} \frac{\tau^1 q^1 N^{1\alpha^1}}{\tau^0 q^0 N^{0\alpha^0}} \right] \qquad (8)$$

The Laspeyres index (CPI) usually calculated by Statistics Canada can be traced from equation (8) under the assumption that between the periods:

(i) The proportion of total direct taxes paid by the median voter remains unchanged $\tau^1 = \tau^0$;

(ii) The unit production cost of public goods remains unchanged: $q^1 = q^0$;

(iii) The goods produced by the State are public goods $\alpha = 0$, or the population of the economy remains unchanged, as does the public goods consumption crowding rate $N^1 = N^0$, $\alpha^1 = \alpha^0$.

In addition to these assumptions, since the traditional CPI is calculated on the basis of net income from direct taxation, $e^0 = p^0(1+t^0)X^0$, and $m^0 = e^0 + T^0$, then (8) unde

these conditions becomes:

$$T^I_{Laspeyres} = 100 \left[\frac{p^1(1+t^1)X^0}{e^0} \left(\frac{e^0}{m^0}\right) + \frac{T^0}{m^0} \right]$$

$$= 100 \left[\frac{C.P.I.}{100} \left(1 - \frac{T^0}{m^0}\right) + \frac{T^0}{m^0} \right]$$

As can be seen, it would be appropriate to make a minor correction to the calculation of the CPI as suggested by this equation. It is clear that $T^I_{Laspeyres}$ < CPI when the CPI > 100. The bias increases if the CPI increases or if $\frac{T^0}{m^0}$ is large!

Indirect and Direct Taxes

One problem associated with the inclusion of public goods in the calculation of a price index is the role of indirect taxes. Several years ago, Kessel [1961] recognized the overstatement of the CPI due to indirect taxes and suggested that they be excluded from the CPI or that direct taxes be included. Paradoxically, however, he wrote (p.523) that the choice of indirect in preference to direct taxes usually produces not only a rise in the prices of public vis-à-vis private goods, but also relative price changes among private goods. Kessel's analysis failed to deal with the consumption of public goods *per se*, an omission that might be at the source of his confusion.

Let us examine the effect on our model of the situation of an increase in the total cost (G) of goods produced by the State. Three possibilities may arise:

(1) The government increases the unit cost (q) of the commodity it produces. This may occur for all government-produced goods subject to a tariff such as the mails, highways, etc. These goods are already included in the calculation of the CPI. Obviously, this case is the most comparable from a theoretical standpoint to that of the privately-produced goods consumed by the median voter.

(2) The government decides it is impossible to estimate q or even to transmit this information to the consumer and decides to increase indirect taxes.[7] Obviously, this decision has the effect of increasing the relative price of private goods and of increasing consumer demand for public goods, and not the inverse as suggested by Kessel.

(3) The same situation described in (2) with respect to the unit cost q also applies here, except that in this case the government decides to modify direct taxes. Assuming it does not modify the redistributive effect of direct taxes, i.e. if τ remains unchanged for the median voter, then, by comparison with (1), only the income effect persists, but by comparison with (2), the misleading signal of relative prices disappears.[8]

If we accept the assumptions that permit the derivation of equation (9), it is easy to see that an indirect tax increase will raise the true Laspeyres index, while an equivalent increase in direct taxes will help to reduce the value of this index. As Kessel [1961] noted, the government's choice of taxation methods makes it difficult to compare the price indexes (as calculated by equation (9)) between countries and between periods. However, this difficulty can be surmounted since the presence of indirect taxes in the calculation of this particular index poses no conceptual problems. The assumption that the elasticity of substitution is null, $\sigma = 0$, yields a utility function of the Leontief type involving fixed proportion commodity consumption by type, i.e. a perfect complementarity between private and public goods within the utility function itself. It then becomes immaterial which of these goods is taxed, just as it would be immaterial whether one taxed the left shoe or the right shoe. Under such circumstances, there is no misdirection and it is still possible for index comparison purposes to convert indirect taxes into their direct tax equivalent or *vice versa* to permit a unique calculation of indexes.

The situation is complicated considerably when the elasticity of substitution σ differs from 0. Not only does the calculation bias in terms of taxation equivalences become more complex, there is considerable doubt as to the economic significance and relevance of the computed and published index. The problem is minimized if the increase in public goods production costs produces an increase in selected indirect taxes on private complements to public goods and permits the substitution of other private goods; in the event of a general indirect tax increase, however, the press release announcing an increase in the CPI on the heels of an indirect tax increase would in fact suggest a relative displacement of demand towards the very public goods whose costs had increased!

Once again, the problem of indirect taxes on the demand for public goods has been ignored in the public choice literature, and it is not the intention of this paper to wage a campaign against the use of indirect taxes to finance government expenditure. The analysis presented here simply points out the conditions under which the inclusion of government-produced goods in the calculation of a cost-of-living index will conform to economic logic. In this context, if the elasticity between private and public goods differs from 0, the notion of the unit production cost of public goods, the proportion of total direct taxes paid by the median voter, the population and the crowding parameter are all relevant dimensions of this analysis, while indirect taxes pose problems.

3. Empirical Considerations

The above discussion has revealed that the elasticity of substitution between private and public goods, σ, and the crowding parameter associated with the consumption of public goods, α, are two important elements in the specification of a cost-of-living index that includes government goods. The estimation of the general index, T^1, of equation (7) requires the specification of the median voter's income and his contribution to total direct taxes, τ, collected by government.

It should be noted that the problem of elasticity of substitution, σ, is more theoretical than practical, as demonstrated by the Lloyd [1975] simulations with a constant utility price index derived from a CES utility function. This fact in no way detracts from the problem of the relevance of the index with the inclusion of indirect taxes following whether σ differs from 0 or not. Moreoever, in the disagregated data assumption, Lloyd [1975] also showed that for a two-level CES function, the question of the value of the elasticity retains its full practical importance.

Theoretically, the crowding parameter can vary between 0 and 1; however, our simulations will be limited to the extreme values associated with the "pure" public goods assumption ($\alpha = 0$) and the "collectively financed private goods" assumption ($\alpha = 1$).[9]

In median voter models, the median voter's income is associated with the median income in the economy as a whole. Moreover, the share of total taxes paid by the median voter is assumed to parallel the ratio of his income to the total income within the economy.

Thus:

$$\tau = \left(\frac{m}{yN}\right) = \left(\frac{m}{y}\right)\frac{1}{N} \tag{9}$$

where $(\frac{m}{y})$ is the ratio of median income to average income, a measure associated with the Gini coefficient of income inequality, under the assumption of a lognormal income distribution law.[10]

The proportionality of the direct tax system in Canada is not above suspicion; however for simulation purposes, this assumption is operational. Moreover, our specification of τ, by equation (4), indicates that the proportion of direct taxes paid by the median voter varies with the amount of indirect tax collected by governments and the budget deficit or surplus situations of these governments. Finally, it would also be possible to define the median voter income level, which differs from the median income level in the economy as a whole. For example, it would be possible to derive the median voter's income from the distribution of incomes over all citizens exercising their right to vote and to subsequently construct the ratio of direct taxes paid by the median voter to the total direct taxes collected in the economy.

The data required to validate such specifications are not easy to obtain. We shall ignore the question of indirect taxes and budget surpluses or deficits but still illustrate the role played by the modification of median voter income on the price index. This will be done by distributing the economy's Gross National Product over the whole of the labour force even though this population excludes pensioners with the right to vote and includes workers or unemployed persons who are too young to vote.[11] Assuming the median voter is a labour force participant and continuing to assume that taxes are proportional to income then:

$$\tau = \frac{m^A}{y^A N^A} = \left(\frac{m^A}{y^A}\right)\frac{1}{N^A} \tag{10}$$

where $y^A = \text{GNP} \div N^A$

N^A = labour force.

The index A denotes labour force participant.

The order of magnitude associated with these assumptions can be compared by using data compiled on the ratio of median income to average income and the other variables involved.[12] In 1980, for example, $\frac{m}{y}$ is equal to 0.76, the total population N is 23.90 million and the labour force N^A is 11,522 million. By equation (9), τ therefore equals 0.000000031 and by equation (10), 0.000000065, so that the relative share of total direct taxes paid by the median voter has doubled with the latter specification.[13] In addition, using our definition of y^A, we see that the median incomes in question for 1980 are in the order of $(\frac{m}{y}) \cdot y = \$10{,}000$ with the total population specification and \$20,000 $(=(\frac{m}{y})y^A)$ with the labour force specification.[14] In principle, we could calculate two indexes that would be associated with these two different definitions of median voter income. In practice, however, we are more concerned with comparing the growth of these two indexes with reference to a common base year, such as 1971 = 100.

The evolution of this variable is one of the advantages of the labour force specification. For example, if the income distribution (as measured by the ratio of median income to average income) is kept constant, an increase in the labour force will reduce the relative share of direct taxes paid by the median voter, since the total population will continue to consume public goods at the actual crowding rate, but the cost of producing these goods will be spread over more citizens.

It should be noted that this specification directly introduces a demographic dimension to the theory of public goods demand via the relative price of these goods. Consequently, this demographic element also enters into the calculation of a constant utility price index incorporating public goods.

Using equation (7), identifying the total income within the economy as the Gross National Product ($GNP = y \cdot N = y^A \cdot N^A$), and taking as a measure of the pure movement

of private prices the Laspeyres index produced by Statistics Canada,[15] the constant utility price index of equation (7) is rewritten:

$$T^{I} = 100 \left[\left(1 - \frac{G^{0}}{P.N.B.^{0}}\right) \left(\frac{\text{Laspeyres}^{1}}{100}\right)^{1-\sigma} \right. \tag{11}$$

$$\left. + \frac{G^{0}}{P.N.B.^{0}} \left(\frac{(\frac{m}{y})^{1} q^{1} N^{1}(\alpha^{1} - 1)}{(\frac{m}{y})^{0} q^{0} N^{0}(\alpha^{0} - 1)} \right)^{1-\sigma} \right]^{\frac{1}{1-\sigma}}$$

since, by budget constraint, $\dfrac{p^{0}(1+t^{0})\,X^{0}}{m^{0}} = 1 - \dfrac{T^{0}}{m^{0}}$

and $\dfrac{T^{0}}{m^{0}} = \dfrac{\tau G^{0}}{m^{0}} = (\dfrac{m}{y})^{0} \dfrac{1}{N^{0}} \dfrac{G^{0}}{m^{0}} = \dfrac{G^{0}}{yN^{0}} = \dfrac{G^{0}}{G.N.P.^{0}}$

Under the assumption associated with equation (10) on the role of the labour force, the following index is obtained using the same substitutions:

$$T^{I}_{A} = 100 \left[\left(1 - \frac{G^{0}}{P.N.B.^{0}}\right)\left(\frac{\text{Laspeyres}^{1}}{100}\right)^{1-\sigma} \right. \tag{12}$$

$$\left. + \frac{G^{0}}{P.N.B.^{0}} \left(\frac{(\frac{m}{y})^{1} q^{1} \; \frac{N^{1}(\alpha^{1})}{N_{A}}}{(\frac{m}{y})^{0} q^{0} \; \frac{N^{0}(\alpha^{0})}{N_{A}}} \right)^{1-\sigma} \right]^{\frac{1}{1-\sigma}}$$

To complete the data set required for our simulations, we estimated the unit production costs of public goods q using the implicit index of current public expenditure on goods and services.[16] Since this index is based on the assumption that public sector productivity remains fixed, it was also alternatively assumed that q can vary at the same rate as the growth of private commodity prices, i.e. the Laspeyres index of the CPI.

Table 1 summarizes the results of our simulations. We quickly note, as Lloyd [1975] did, that for all intents and purposes, the indexes based on a CES utility function do not vary with the value of the elasticity of substitution between private and public goods. However, the indexes specified with the labour force constitute an exception in cases where the unit production cost of public goods varies at the same rate as the Laspeyres consumer price index ($q = L$). In such cases, the computed indexes decline gradually and steadily when the elasticity of substitution rises to an inter-index cumulative growth rate discrepancy level of 3% for the period 1971-1980.

As expected, the level of the indexes is lower after 1971 in cases where the government-produced goods are pure public goods ($\alpha = 0$) as opposed to collectively financed private goods ($\alpha = 1$). The size of the differences in terms of index growth for the period 1971-1980 is about 10% with the labour force specification as compared to about 8% with the total population specification.

The indexes are also highly sensitive to the unit production cost (q) specification of government goods. When the implicit price index of public expenditure on goods and services is used to estimate q, rather than the Laspeyres consumer price index ($q = L$), the level of the first index is far higher after 1971, resulting in cumulative growth rate discrepancies in the order of 15% for the period 1971-1980.

It should be noted that during this period, the CPI rose 110.6% (to 210.6 in 1980) as compared with 154.4% (to 254.4 in 1980) for the implicit price index of public expenditure. Should the true cost index (q) fall somewhere in between these two measures, it is almost certain to fall closer to the latter. In labour alone, for example, we know that government employees have generally received cost-of-living increases in response to changes in the CPI and their increasing union strengh which, in recent years, has helped them enjoy higher

TABLE 1. Simulations of Constant Utility Price Indexes Including Goods Produced by Governments (1948-1980: selected years)

	Total population specification								Labour force specification							
	$\sigma = 0$				$\sigma = 1.5$				$\sigma = 0$				$\sigma = 1.5$			
	$\alpha = 0$		$\alpha = 1$		$\alpha = 0$		$\alpha = 1$		$\alpha = 0$		$\alpha = 1$		$\alpha = 0$		$\alpha = 1$	
	q = q	q = L	q = q	q = L	q = q	q = L	q = q	q = L	q = q	q = L	q = q	q = L	q = q	q = L	q = q	q = L
1948	57.50	73.51	48.58	58.10	57.44	68.57	47.02	57.98	58.13	74.60	48.96	58.75	58.01	69.14	47.57	58.56
1950	60.81	76.00	52.56	62.22	60.79	72.05	51.04	62.12	62.23	78.38	53.34	63.73	62.13	73.38	52.32	63.45
1955	68.43	80.18	61.49	70.04	68.41	77.77	60.67	69.92	71.48	84.63	63.71	73.29	71.20	80.55	63.40	72.71
1960	75.07	83.32	70.30	77.13	75.06	82.11	69.99	77.00	78.28	87.48	72.96	80.58	78.02	85.08	72.93	79.97
1965	80.42	85.68	77.81	82.61	80.42	85.29	77.68	82.54	83.36	89.16	80.49	85.77	83.23	88.11	80.49	85.36
1970	97.02	98.14	96.57	97.68	97.02	98.13	96.57	97.68	97.64	98.78	97.19	98.31	97.64	98.75	97.19	98.30
1971	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
1972	105.26	104.43	105.71	104.87	105.25	104.43	105.70	104.87	104.56	103.75	105.01	104.18	104.56	103.74	105.01	104.18
1973	112.82	111.93	113.82	112.91	112.82	111.92	113.81	112.91	110.88	110.03	111.84	110.97	110.84	109.94	111.83	110.93
1974	124.89	123.55	126.73	125.34	124.89	123.53	126.70	125.34	121.69	120.45	123.41	122.11	121.57	120.22	123.38	122.02
1975	142.86	136.72	145.70	139.24	142.69	136.69	145.25	139.24	137.52	131.98	140.08	134.25	137.51	131.54	140.06	134.07
1976	155.31	145.78	159.31	149.15	154.97	145.69	158.45	149.15	148.80	140.30	152.37	143.30	148.80	139.58	152.27	143.01
1977	167.72	156.61	172.91	160.92	167.35	156.46	171.84	160.92	159.71	149.96	164.27	153.74	159.70	148.88	164.17	153.31
1978	179.37	168.25	185.34	173.24	179.24	167.86	184.64	173.22	170.06	160.46	175.21	164.77	169.85	158.58	175.21	163.87
1979	195.17	183.52	202.32	289.53	195.07	183.02	201.54	189.51	183.77	173.93	189.81	179.01	183.34	171.52	189.80	177.87
1980	215.65	201.59	224.50	208.92	215.50	201.05	223.41	208.90	201.79	190.12	204.14	196.20	201.27	187.00	209.13	194.74

Source: S1039, S1040, 1013 and 1022 computer printouts.

salary increases than their private sector counterparts.[17]

Finally, we note that the index constructed with the labour force specification has grown at a slower rate than that based on the total population specification, again resulting in growth rate discrepancies in the order of 15%. This is due to the discrepancy between the growth rates of Canada's labour force and its total population between 1971 and 1980 (33.5% as compared to 10.8%). This spectacular growth of the labour force coupled with the comparatively less impressive growth of the total population has lowered the price of public goods for the median voter.

Table 2 summarizes the main results obtained from the simulations in the form of rates of increase of the cost of living.

As Table 2 indicates, the variability in the simulation results is reflected in a wide variation in the inflation rates and in the cost-of-living figures calculated using the various plausible indexes.

It is possible to observe the order of magnitude of the differences between the indexes obtained from these simulations and the usual ones or to compare them with the changes associated with other simulations or experiments suggested by price index theory. In particular, it will be recalled that the Statistics Canada *Consumer Price Index Reference Paper* (1982, pp.93-94) shows that for the overall index, the cumulative three-year discrepancy between the index based on 1974 goods and that derived from an update of the basket of 1978 goods was only 0.5% in relative terms, for an average of slightly more than 0.1% per year.

The results of Table 2 show that the implicit public expenditure index generated the largest cost-of-living increases, followed by the CPI. The indexes including public and private goods have rates comparable to or greater than the CPI under the total population assumption of the relative share of total direct taxes paid by the median voter; these rates are comparable to or less than the CPI under the labour force assumption. In the latter case, the index most comparable to the CPI in practice, if evidently not in theory, implies the implicit price index of public spending, q, and a public goods consumption crowding

TABLE 2. Annual Rate of Increase (in percent) of the Cost of Living or Inflation Rate Using Different Indexes

	Total population specification $\sigma = 1.5$				Labour force specification $\sigma = 1.5$				Implicit price index of public expenditure	Laspeyres Consumer Price Index (CPI)
	$\alpha = 0$		$\alpha = 1$		$\alpha = 0$		$\alpha = 1$			
Period	q=q	q=L	q=q	q=L	q=q	q=L	q=q	q=L	q	C.P.I.
1950-55	2.5	1.6	3.7	2.5	2.9	1.9	4.2	2.9	5.7	2.5
1955-60	1.9	1.1	3.1	2.0	1.9	1.1	3.0	2.0	4.8	2.0
1960-65	1.4	0.8	2.2	1.4	1.3	0.7	2.1	1.3	3.8	1.7
1965-70	4.1	3.0	4.9	3.7	3.5	2.4	4.1	3.0	7.6	4.1
1970-75	9.4	7.8	10.1	8.5	8.2	6.6	8.8	7.3	13.1	8.5
1975-80	10.2	9.4	10.8	10.0	9.2	8.4	9.9	9.0	12.6	10.4

level equal to unity ($\alpha = 1$).

Obviously, the simulations leave unanswered the question of which price index is closest to economic reality. However, we believe the approach proposed in this paper permits an adequate response to this question in providing sufficient elements for the construction of the median voter's demand for public goods. This in turn permits the definition of various specifications and parameter values which more closely resemble statistical reality.

Maximizing the Lagrangian formed from the utility function of equation (6) and the budget constraint resulting from a substitution of equations (2) and (4) in (3), yields (from the equilibrium conditions) the following expression:

$$\frac{X}{H} = \left[\frac{(1-\delta)}{\delta} \; \frac{p(1+t)}{\tau \, q N^{\alpha}} \right]^{\frac{1}{\eta - 1}} \tag{13}$$

Multiplying this equation by $\dfrac{p(1+\tau)}{\tau q N^{\alpha}}$, and reusing equation (2) and the budget constraint specified above yields:

$$\frac{(m-\tau G)}{\tau G} = \frac{(1-\delta)}{\delta}^{\frac{1}{\eta - 1}} \left[\frac{p(1+t)}{\tau \, q N^{\alpha}} \right]^{\frac{\eta}{\eta - 1}} \tag{14}$$

A statistically estimable expression is easily obtained by specifying according to equation (9) or (10). For example, using the labour force specification of τ and substituting it in expression (14), after a few rearrangements and the application of the logarithmic function, yields:

$$\ln \left(\frac{1 - \dfrac{G}{P.N.B.}}{\dfrac{G}{P.N.B.}} \right) = \beta_0 + \beta_1 \ln \left(\frac{p(1+t) \cdot N^{A}}{(\frac{m}{y})\, q} \right) + \beta_1 \beta_2 \log N + \epsilon \tag{15}$$

where $\beta_0 = \frac{1}{\eta - 1} \ln\left(\frac{(1-\delta)}{\delta}\right).$

$\beta_1 = \frac{\eta}{\eta - 1}$ so that $1 - \beta_1 = \sigma$: the elasticity of substitution.

$\beta_2 = -\alpha$, where α is the crowding parameter.

ϵ = a random term with the usual properties, added on an ad hoc basis.

The data of the model covering the period 1947-1980 were used to construct an exploratory estimate of equation (15) and the equation corresponding to the total population specification of τ. Several difficulties arose at the econometric level, particularly that of obtaining parameter estimates compatible with our theoretical specifications (for example, $0 \leq \alpha \leq 1$). The major problem of multicolinearity between variables should also be noted. Only the results reported in Table 3, obtained for the labour force specification using specific values of α, appear suitable.

Based on marginal statistical evidence and in light of the value of the parameter δ, the most acceptable regression is that involving the specification of government goods as collectively financed private goods ($\alpha = 1$). It should be noted that in studies dealing with the demand for government goods, the crowding parameter in the consumption of public goods has always been estimated at unity.[18] These econometric results also suggest an elasticity of substitution between private and public goods different from 0 and close to 1.5

In comparison with our simulations, these results would lend weight to the labour force specification of the price index, the unit production cost q and a crowding parameter equal to 1. From a practical standpoint, the latter index has followed the usual CPI relatively closely in recent years. However, the regression analysis confirms that this similarity of movement stems from numerical coincidence.

TABLE 3. **Estimating Equation (15) by OLSQ, Corrected for Autocorrelation in the Residuals**

Assumption	Regression parameters		Structural parameters		$\bar{R}^2$	Likelihood function logarithm
	β_0	β_1	δ	σ		
0	7.87 (6.07)	-0.46 (0.37)	0.99	1.46	0.963	52.73
1	-0.159 (0.38)	-0.616 a (0.36)	0.476	1.61	0.964	53.36

( ): standard deviation.
a: significant at 0.90, two-tail test.
Source: Computer printouts S1046, S1215.

4. Conclusion

The median voter model has permitted the integration of public goods into the construction of a constant utility price index within a theoretical framework.

Moreover, at an aggregate level, we have demonstrated the possibility of defining an empirical counterpart to this index. While the resulting index (based on the data used) has produced a movement similar to that of the CPI in recent years, this result would appear to be a mere numerical coincidence and there is no evidence to indicate it is anything else.

The index as defined still has several shortcomings. For example, the problem of indirect taxes raises the question of the index's relevance from an economic logic standpoint. In addition, the index's level and movement through time vary widely with the value of certain parameters, particularly that of the proportion of total direct taxes paid by the median voter. Finally, the operational effectiveness of the index at a disaggregated level is unknown.

However, the approach is theoretically solid and could be used in other studies of price indexes and government-produced goods. For example, it would be possible to examine the impact on the index of the transition of the production of a commodity from the private sector to the public sector, as happened in the health sector a few years ago. In the traditional CPI, this commodity would be removed from the index and if its price was greater than average, the index would fall; if less than average, the index would rise. More importantly, we know the index is less representative of price movements in the economy as a whole. With our model we believe it would be possible to modify the production function of public commodity H (equation (2)) and to examine this problem as a question of economies of scale and production by sector.[19]

Footnotes

[1] Gillingham and Greenlees [1983] discuss the incorporation of direct and indirect taxes in the CPI on the basis of an income-defined COL index in opposition to an expenditure-defined COL index. Their procedure measures the before-tax income which is required to attain a given level of satisfaction (in the CPI sense) given the evolution of the taxes. This approach leaves out the dimension "consumption of government goods" that those taxes permit.

[2] See Aranson and Ordeshook [1981] for a general critique.

[3] The bureaucratic and "logrolling" models are often presented as alternatives to the median voter model. These models are also controversial and ill-suited to a discussion on price index theory. One exception of note, however, is the Bös [1978] study which uses a utility function of the politicians to link the cost-of-living indexes to a public pricing policy.

[4] To our knowledge, the literature on median voter models contains no discussion of indirect taxes and budget deficits or surpluses.

[5] $\tau(\ ) = \tau$ to simplify the notation.

[6] See Lloyd [1975, p.305], equation 14, for a specific demonstration of this property.

[7] Certain levels of government are permitted to level only indirect taxes.

[8] Obviously, the price effect could resurface if the government were to modify the proportion of total direct taxes paid by the median voter, i.e. τ. Evidently, these factors have a collective impact on the demand for public goods and on the relative share of government spending in the economy as a whole.

[9] The latest expression is that of Aranson and Ordeshook [1981].

[0] See Aitchison and Brown [1963].

[1] The labour force does not include inmates, members of the Armed Forces, Indians living on reserves, or residents of the Yukon and the Northwest Territories. Source: *Canadian Statistical Review*, 11-003E, Section 4, Table 3.

[2] Source: $\frac{m}{y}$: estimated from Statistics Canada data, Catalogue 13-559, by interpolating certain data for the period 1951-1973 and through extrapolation for 1948-1951 and 1974-1980 by regressing the SC data on the previously transformed Gini coefficients of Buse's [1980] study. N: SC, 13-001.

[3] By assumption $\left(\frac{m^A}{y^A}\right) = \left(\frac{m}{y}\right)$.

Source: Y = GNP: SC, 13-001.

Source: Laspeyres CPI: *Canadian Statistical Review*, SC, 13-001.

Source: G: Statistics Canada, Catalogue 13-001.
q: 1947-1974, Statistics Canada, Catalogue 13-531;
1974-1980, Statistics Canada, Catalogue 13-001.

[17] On the growth of wages in the public and private sectors, see the studies of Lacroix and Cousineau [1977] and Lacroix [1982].

[18] One exception is the Dudley-Montmarquette [1981] study on the estimation of the demand for military spending.

[19] In a recent paper, Dudley and Montmarquette [1983] specified a median voter model of the demand for public goods using a more general production function of H than the one defined by equation (2).

References

Aitchison, J. and J.A.C. Brown, *The Lognormal Distribution with Special Reference to its Use in Economics*, Cambridge University Press, 1963.

Aranson, P.H. and P.C. Ordeshook, "Regulation, Redistribution and Public Choice", *Public Choice*, 37, 1981, pp.69-100.

Bergstrom, T.C. and R.P. Goodman, "Private Demand for Public Goods", *American Economic Review*, 63, 1973, pp.280-296.

Borcherding, T.E. and R.T. Deacon, "The Demand for Services of Non-Federal Governments", *American Economic Review*, 62, 1972, pp.891-901.

Bös, D., "Cost of Living Indices and Public Pricing", *Economica*, 45, 1978, pp.69-79.

Bowen, H.R., "The Interpretation of Voting in the Allocation of Economic Resources", *Quarterly Journal of Economics*, 58, 1943, pp.27-48.

Buse, A., "The Cyclical Behavior of the Size Distribution of Income in Canada", Discussion Paper No. 80-13, University of Alberta, 1980.

Dudley, L. and C. Montmarquette, "The Effects of Non-clearing Labor Markets on the Demand for Public Spending", Unpublished paper, January 1983.

"The Demand for Military Expenditures: An International Comparison", *Public Choice*, 37, 1981, pp.5-31.

Gillingham, R. et J.S. Greenlees, "The Incorporation of Direct Taxes into a Consumer Price Index", *Price Level Measurement: Proceedings From a Conference Sponsored by Statistics Canada*, 1983.

Goldberger, A.S., "Functional Form and Utility: A Review of Consumer Demand Theory", University of Wisconsin: Systems Formulation, Methodology and Policy Workshop Paper 6703, October 1967.

Hotelling, H., "Stability in Competition", *Economic Journal*, 39, 1929, pp.41-57.

Kessel, R.A., "The Measurement and Economics Implications of the Inclusion of Indirect Taxes in the Consumer Price Index", in *The Price Statistics of the Federal Government*, N.B.E.R., 1961, Chapter 12.

Klein, L.R. and R. Rubin, "A Constant-Utility Index of the Cost of Living", *The Review of Economic Studies*, XV, 1948-49, pp.84-87.

Lacroix, R., *Les conséquences des disparités de salaires entre secteurs privé et public au Québec*, Paper No. 16, Conseil du patronat du Québec, March 1982.

__________ and J.M. Cousineau, *La détermination des salaires dans le monde des grandes conventions collectives: une analyse des secteurs privé et public*, Economic Council of Canada, Ottawa, 1977.

Lloyd, P.J., "Substitution Effects and Biases in Non-True Price Indices", *American Economic Review*, 65 (3), June 1975, pp.301-313.

Statistics Canada, *The Consumer Price Index Reference Paper*, Catalogue 62-553, Occasional, May 1982, 151 pages.

LES BIENS PUBLICS ET LES INDICES DE PRIX

To provide you with a version in the official language of your choice, the French text is preceeded by the English text (p.655) in this publication.

Claude Montmarquette
Université de Montréal

RÉSUMÉ

L'objectif du texte est d'inclure les biens publics dans le calcul d'un indice de prix à utilité constante couvrant tous les prix des biens de l'économie. Nous montrons comment le modèle du voteur médian, emprunté à la littérature des choix publics, permet de relier les biens publics à la littérature traditionnelle sur la mesure des prix. Diverses simulations et résultats empiriques sont présentés, utilisant des données canadiennes couvrant la période de 1948 à 1980.

L'indice des prix à la consommation, comme l'a rappelé récemment, avec insistance, un document de référence publié par Statistique Canada (1982, p. 9), ne vise qu'à mesurer le mouvement pur des prix. Plus explicitement, l'I.P.C. mesure les changements de prix en comparant, dans le temps, le coût d'un panier de biens de quantité et de qualité invariables ou équivalentes et déterminés en fonction des achats effectués par une population-cible. Sans même discuter des difficultés inhérentes au calcul de l'I.P.C., l'insistance de Statistique Canada provient de ce que l'utilisation de données agrégées devient rapidement abusive faute de présenter aux utilisateurs des alternatives valables.

L'utilisation de l'I.P.C. en tant que mesure du coût de la vie pour fins d'indexation des loyers, des actifs ou autres biens et services, des salaires et autres revenus a reçu beaucoup d'attention de la part des chercheurs. Leurs études ont visé à définir les limites, biais ou conditions de cet usage. Les économistes, entre autres, ont souvent mis en cause l'arbitraire même de cette mesure des mouvements des prix étant donné le confinement de son application à un groupe spécifique d'individus et à un groupe donné de biens et services.

L'idée d'une mise à jour de l'I.P.C., fondé sur des dépenses plus récentes, le problème des changements de qualité dans les biens, l'utilisation des indices spatiaux et de la théorie des sous-indices, pour ne citer que quelques champs de préoccupation, sont des réponses partielles à ces critiques. Une lacune importante demeure cependant, concernant le rôle des biens publics dans la production d'un indice de prix. La littérature est virtuellement demeurée muette sur cette question depuis la discussion Kessel [1961] sur le rôle des taxes directes et indirectes dans le calcul de l'I.P.C.[1] Globalement, près de 40% des dépenses totales de l'économie sont effectuées par les gouvernements, et l'I.P.C. actuellement calculé n'est concerné que par une très faible portion de ces dépenses. Un concept d'indice de prix le moindrement pertinent se doit de traiter adéquatement des biens publics.

L'objectif du présent texte est d'inclure les biens publics dans le calcul d'un indice de prix à utilité constante couvrant tous les prix des biens de l'économie.

Dans une première section de l'étude, nous montrons comment le modèle du voteur médian, emprunté à la littérature des choix publics, permet de relier les biens publics à la littérature traditionnelle sur la mesure des prix.

Dans une deuxième section, nous définissons un indice de prix à utilité constante incluant les biens produits par les gouvernements. Nous y présentons, d'un point de vue abstrait, certaines dimensions de cet indice témoignant de son intérêt et de sa pertinence en plus de le comparer aux indices d'utilisation courante. Les problèmes associés aux taxes indirectes et directes dans la définition de l'indice sont également discutés dans cette section.

Dans la troisième section, diverses simulations et résultats empiriques sont présentés, utilisant des données canadiennes couvrant la période de 1948 à 1980.

Enfin, le texte se termine en soulignant la pertinence théorique de l'indice proposé et son opérationnalité au plan empirique. Plus généralement, l'approche retenue dans cette étude ouvre de nouveaux horizons dans le traitement de questions liées aux indices de prix et à la production du secteur public.

1. Le modèle du voteur médian

L'inclusion des biens gouvernementaux dans le calcul d'un indice de prix à la consommation présente la difficulté de distinguer entre quantité et prix. En général, les services et biens des gouvernements ne sont pas tarifiés et la nature même des biens publics, donc susceptibles d'être consommés par tous, n'est pas pour en simplifier la mesure quantitative. Par ailleurs, pour que l'indice de prix soit une véritable mesure du coût de la vie, il faut inclure tous les biens d'une économie dans une fonction de préférence. Dans le cadre des biens publics, le problème traditionnel du resquilleur (free riders) complique cette analyse en termes de fonction de préférence.

Le modèle du voteur médian, que l'on retrouve dans la littérature sur la théorie des choix publics (voir, par exemple, Bergstrom et Goodman [1973], Borcherding et Deacon [1972], Dudley et Montmarquette [1981], offre une porte de sortie intéressante pour pallier à ces difficultés. En effet, sous certaines conditions, entre autres celle des préférences symétriques des voteurs autour d'un événement électoral unique, le modèle du voteur médian, tel qu'initialement formulé par Hotelling [1929] et Bowen [1943], démontre que ce sont les préférences de ce voteur qui déterminent les choix publics. Il s'ensuit, comme l'exprime très bien Borcherding et Deacon [1972, p.892], qu'*ex post*, nous sommes alors en mesure d'observer les préférences *ex ante* du voteur médian, ce qui nous libère du problème d'observer les préférences d'un resquilleur.

Dans la théorie des choix publics, le modèle du voteur médian ne fait pas nécessairement l'unanimité; pour le moment, les avis sont partagés sur sa valeur et sa pertinence.[2] On peut néanmoins accepter que ce modèle puisse servir de représentation valable de la réalité. De plus, pour la théorie des indices de prix, le modèle du voteur médian présente cet avantage de traiter tous les biens de la même façon.[3]

Considérons, en effet, la fonction d'utilité suivante du voteur médian:

$$U\,(X,H) \qquad (1)$$

où X est un bien privé composé et H, un bien produit par le gouvernement. Soit:

$$H = \frac{G}{qN^{\alpha}} \qquad (2)$$

où G représente les dépenses gouvernementales totales. L'unité de mesure de H est définie de telle sorte que son coût unitaire est égal à q. N représente la population et α un paramètre d'encombrement. Si $\alpha = 0$, alors H est un bien public pur: le voteur médian et chacun des individus dans l'économie consomment la production totale H. Si α est égal à l'unité ($\alpha = 1$), H est l'équivalent d'un bien privé dans la mesure où chaque individu, dans l'économie, ne profite que de la $\frac{1}{N}$ ième partie de la production totale H du gouvernement.

La contrainte budgétaire du voteur médian se définit comme:

$$pX + tpX + T = m \qquad (3)$$

où p est le prix unitaire du bien privé X; t est le taux de taxe indirecte en vigueur (de sorte que tpX est la valeur des taxes indirectes payées par le voteur médian) T est la valeur des taxes directes payées par le voteur médian; m, le revenu du voteur médian.

Établissons l'existence d'une relation entre les taxes directes payées par le voteur médian et les dépenses gouvernementales:

$$T = \tau(\,)\,G \qquad (4)$$

τ représente la part des taxes directes totales payée par le voteur médian. Cette part **est fonction** de la structure de taxation des gouvernements, du montant des dépenses gouvernementales financées par les taxes indirectes et des surplus ou déficits budgétaires.[4]

Substituons (2) et (4) dans (3) et maximisons le Lagrangien formé de la fonction d'utilité (1) et de la contrainte budgétaire telle que redéfinie. Nous obtenons, en plus de cette contrainte, les conditions usuelles d'équilibre qui, exprimées sous forme de rapport des utilités

marginales U_X et U_H, donnent l'équation suivante[5]:

$$\frac{U_X}{U_H} = \frac{p\,(1 + t)}{\tau q N^{\alpha}} \qquad (5)$$

A l'équilibre, le rapport des utilités marginales est égal au rapport des prix: il est donc clair que l'expression $\tau q N^{\alpha}$ identifie le prix de H pour le voteur médian. On notera que ce prix dépend de la part des taxes directes totales payée par le voteur médian, alors que les taxes indirectes ne touchent que le prix des biens privés: en d'autres termes, toutes autres choses étant égales par ailleurs, une hausse des taxes indirectes va réduire paradoxalement le prix relatif des biens publics et donc, théoriquement, accroître la demande pour ces biens publics. Ce point semble avoir échappé aux analystes de la théorie des choix publics. Par ailleurs et pour reprendre les termes de Bergstrom et Goodman (1973, p. 282), si α est près de zéro, il y a des économies substantielles pour une économie avec une forte population, puisque plus de consommateurs peuvent partager les coûts des biens publics sans trop souffrir d'effets d'encombrement. D'autre part, si α tourne autour de l'unité, les gains dus au partage des coûts entre les consommateurs sont contrebalancés par la nécessité de devoir partager les biens gouvernementaux entre ces mêmes personnes.

À l'aide de ce modèle, il devient alors possible d'utiliser diverses procédures standards, comme l'intégration des fonctions de demande à la Klein et Rubin [1948-1949], l'approche par la fonction d'utilité indirecte de Goldberger [1967] ou une fonction de dépense telle qu'appliquée par Lloyd [1975] pour définir un indice du coût de la vie à utilité constante. Cet indice, rappelons-le, correspond au ratio entre le revenu maintenant nécessaire et le revenu de la période de référence qui laisse le consommateur sur une même courbe d'indifférence malgré des changements de prix survenus entre les périodes.

2. Indice de prix incluant les biens publics

Spécifions la fonction d'utilité (1) comme une fonction C.E.S.:

$$U = [(1-\delta)H^{\eta} + \delta X^{\eta}]^{1/\eta} \qquad (6)$$

où δ est un paramètre; l'élasticité de substitution entre biens privés et biens publics est définie comme $\sigma = \frac{1}{1-\eta}$. Retenant la contrainte budgétaire du voteur médian et utilisant cette fonction d'utilité C.E.S. qui inclut les biens publics, une application des fonctions de dépenses de Lloyd [1975] nous définit un indice de prix à utilité constante entre deux périodes:

$$T^I = 100 \left[\frac{p^0(1+t^0)X^0}{m^0} \left(\frac{p^1(1+t^1)}{p^0(1+t^0)} \right)^{1-\sigma} + \frac{T^0}{m^0} \left(\frac{\tau^1 q^1 N^{1\,\alpha^1}}{\tau^0 q^0 N^{0\,\alpha^0}} \right)^{1-\sigma} \right]^{\frac{1}{1-\sigma}} \qquad (7)$$

où les indices ^{0} et ^{1} identifient respectivement la période de base et une autre période; T^0 correspond à la part de taxes directes payée par le voteur médian, puisque le prix du bien public multiplié par la qualité demandée donne ces taxes: $\tau q N^{\alpha} \cdot H = \tau G$ par l'équation (2), ce qui finalement est égal à T par l'équation (4); m^0 représente le revenu nominal brut du voteur médian de la période de référence; σ est l'élasticité partielle et constante de substitution entre les biens privés et les biens publics.

Comme l'a souligné Lloyd [1975] dans le contexte d'une fonction homothétique comme la C.E.S., l'indice de prix à utilité constante est indépendant du niveau initial du revenu, puisque la part de chaque bien dans les dépenses totales dépend essentiellement du niveau des prix relatifs.[6] L'indice est une moyenne pondérée d'ordre $(1-\sigma)$ dans les prix des biens, les poids représentant les parts respectives dépensées sur chaque bien au cours de la période de référence.

Un des avantages de cet indice, également noté par Lloyd, est le cas particulier où $\sigma = 0$: l'indice devient alors une moyenne arithmétique pondérée, i.e. un indice Laspeyres.

$$T^I_{Laspeyres} = 100 \left[\frac{p^1\,(1+t^1)X^0}{m^0} + \frac{T^0}{m^0} \frac{\tau^1 q^1 N^{1\,\alpha^1}}{\tau^0 q^0 N^{0\,\alpha^0}} . \right] \qquad (8)$$

On peut retracer, de l'équation (8), l'indice Laspeyres (I.P.C.) usuellement calculé par Statistique Canada, en supposant qu'entre les périodes:

i) la part des taxes directes totales payée par le voteur médian demeure inchangée: $\tau^1 = \tau^0$;

ii) aucun changement ne survient dans le coût unitaire de production des biens publics: $q^1 = q^0$;

iii) les biens produits par l'État sont des biens publics $\alpha = 0$, ou que la population de l'économie reste inchangée et que l'on observe le même degré d'encombrement de consommation des biens publics: $N^1 = N^0$, $\alpha^1 = \alpha^0$.

En plus de ces hypothèses, comme l'I.P.C. traditionnel se calcule sur la base du revenu net d'impôt direct, $e^0 = p^0(1 + t^0)X^0$, et que $m^0 = e^0 + T^0$, alors (8) sous ces conditions devient:

$$T^I_{\text{Laspeyres}} = 100 \left[\frac{p^1(1+t^1)X^0}{e^0} \left(\frac{e^0}{m^0}\right) + \frac{T^0}{m^0} \right]$$

$$= 100 \left[\frac{\text{I.P.C.}}{100} \left(1 - \frac{T^0}{m^0}\right) + \frac{T^0}{m^0} \right]$$

On voit qu'une correction minimale à apporter au calcul de l'I.P.C. serait celle suggérée par cette équation. Il est clair que $T^I_{\text{Laspeyres}}$ < I.P.C. lorsque I.P.C. > 100. Le biais est d'autant plus grand que l'I.P.C. est élevé ou que $\frac{T^0}{m^0}$ est important!

Les taxes indirectes et directes

Une des difficultés de l'inclusion des biens publics dans la détermination d'un indice de prix concerne le rôle des taxes indirectes. Il y a plusieurs années, Kessel [1961] soulignait dans le calcul de l'I.P.C., le biais vers le haut dû aux taxes indirectes et suggérait alors

l'exclusion des taxes indirectes de l'I.P.C. ou l'inclusion de taxes directes. Paradoxalement, cependant, il écrivit (p. 523) ''que, généralement, le choix des taxes indirectes de préférence aux taxes directes produit, non seulement une hausse des prix des biens publics vis-à-vis les biens privés, mais aussi des changements de prix relatifs entre les biens privés''[7]. L'analyse de Kessel ne concernait pas comme telle la demande des biens publics, ce qui pourrait être à la source de la confusion.

Reprenons, dans le cadre de notre modèle, la situation d'une hausse du coût total(G) des biens produits par l'État. Trois possibilités peuvent surgir:

1) Le gouvernement hausse le coût unitaire (q) du bien qu'il produit. Ce cas est possible pour tous les produits tarifiés du gouvernement, comme par exemple: les postes, les autoroutes... Ces biens font d'ailleurs déjà partie du calcul de l'I.P.C. et ce cas est évidemment le plus théoriquement comparable aux biens privés consommés par le voteur médian.

2) Le gouvernement juge qu'il lui est impossible d'estimer q ou même simplement de transmettre cette information au consommateur et décide une hausse des taxes indirectes.[8] Ce résultat a évidemment pour conséquence de hausser le prix relatif des biens privés, en plus d'amener le consommateur à demander davantage de biens publics et non l'inverse comme l'a prétendu Kessel.

3) La même situation décrite en 2) par rapport au coût unitaire, q, s'applique, mais ici, le gouvernement choisit de modifier les taxes directes. En supposant qu'il ne modifie pas l'incidence redistributive des taxes directes, c'est-à-dire si τ demeure inchangé pour le voteur médian, alors, par rapport à 1), seul l'effet revenu persiste, mais, par rapport à 2), le faux signal des prix relatifs disparaît.[9]

Si nous acceptons les hypothèses qui nous permettent de dériver l'équation (9), il est facile de voir qu'une hausse des taxes indirectes va hausser le vrai indice Laspeyres, alors qu'une hausse équivalente des taxes directes contribuera à diminuer la valeur de cet indice. Comme l'avait fait remarquer Kessel [1961], selon que les gouvernements optent pour l'une ou l'autre forme de taxation, les indices de prix, tels que calculés par l'équation (9), sont difficilement comparables entre pays et entre périodes. Cette difficulté est surmontable,

cependant, du fait que la présence des taxes indirectes dans le calcul de cet indice particulier ne présente pas un embarras conceptuel. En effet, dans l'hypothèse d'une élasticité de substitution nulle, $\sigma = 0$, nous obtenons une fonction d'utilité de type Leontief impliquant une consommation à proportion fixe entre biens, c'est-à-dire une complémentarité parfaite entre biens privés et biens publics, se situant au niveau même de la fonction d'utilité. Il devient alors indifférent de taxer l'un ou l'autre des biens comme c'est le cas d'une taxe sur le soulier droit ou sur le soulier gauche. Dans ces circonstances, il n'y a pas de signal pervers et il est toujours possible, pour fins de comparaisons entre indices, de traduire les taxes indirectes en montants équivalents de taxes directes ou inversement, pour être en mesure de procéder à un calcul unique d'indices.

Par ailleurs, là où la situation se complique de façon considérable, c'est lorsque l'élasticité de substitution, σ, diffère de zéro. Non seulement le biais de calcul en termes de montants équivalents de taxes devient-il plus complexe, mais il faut sérieusement se demander si l'indice calculé et diffusé dans le public conserve une signification et une pertinence économique. Le problème est minimisé si la hausse des coûts de production des biens publics entraîne une hausse de taxes indirectes sélectives sur des biens privés complémentaires aux biens publics de façon à permettre des substitutions par d'autres biens privés; mais dans le cas d'une hause généralisée des taxes indirectes, le communiqué de presse qui annoncerait une hausse de l'I.P.C. suite à une hausse des taxes indirectes suggérerait un déplacement relatif de la demande vers des biens publics dont les coûts se sont haussés!

Encore une fois, le problème des taxes indirectes sur la demande de biens publics a été ignoré dans la littérature des choix publics et l'objectif de cet article n'est pas de lancer une campagne contre les recours aux taxes indirectes pour financer les dépenses gouvernementales. L'analyse présentée ici nous signale à quelles conditions l'inclusion des biens produits par les gouvernements dans le calcul d'un indice du coût de la vie est conforme à la logique économique. A cet égard, si l'élasticité entre biens publics diffère de zéro, la notion de coût unitaire de production des biens publics, la part des taxes directes totales payée par le voteur médian, la population et le paramètre d'encombrement sont des dimensions pertinentes, alors que les taxes indirectes présentent des difficultés.

3. Considérations empiriques

Il ressort des discussions précédentes que deux des éléments importants quant à la spécification d'un indice du coût de la vie incluant les biens gouvernementaux concernent l'élasticité de substitution entre biens privés et biens publics, σ, et le paramètre d'encombrement, α, associé à la consommation de biens publics. D'autres part, pour estimer l'indice général, T^I, de l'équation (7), il nous faut spécifier le revenu du voteur médian et sa contribution aux taxes directes totales, τ, perçues par les gouvernements.

Notons immédiatement que le problème de l'élasticité de substitution, τ, est plus théorique que pratique comme l'ont démontré les simulations de Lloyd [1975] avec un indice de prix à utilité constante tiré d'une fonction d'utilité C.E.S. Ce poin' n'enlève rien, cependant, à la pertinence de l'indice relativement à l'inclusion des taxes indirectes, selon que σ diffère de zéro ou non. De plus, dans l'hypothèse de données désagrégées, Lloyd [1975] a également démontré que, pour une fonction C.E.S. à deux niveaux, la question de la valeur de l'élasticité garde toute son importance pratique.

Le paramètre d'encombrement peut théoriquement varier entre 0 et 1, mais pour fins de simulations nous n'allons retenir que les valeurs extrêmes associées à l'hypothèse du bien public "pur", ($\alpha = 0$) et celles des "biens privés collectivement financés" ($\alpha = 1$).[10]

Dans le modèle du voteur médian, le revenu de celui-ci est associé au revenu médian dans la distribution des revenus dans l'économie. De plus, on suppose que la part des taxes totales payée par le voteur médian est proportionnelle à son revenu par rapport à la somme totale des revenus dans l'économie.

$$\tau = \left(\frac{m}{yN}\right) = \left(\frac{m}{y}\right)\frac{1}{N} \tag{9}$$

où $(\frac{m}{y})$ est le rapport du revenu médian sur le revenu moyen, une mesure associée au coefficient Gini d'inégalité des revenus, sous l'hypothèse d'une loi lognormale de distribution des revenus.[11]

Au Canada, la proportionnalité du système de taxes directes n'est pas évidente, mais, pour fins de simulations, cette hypothèse est opérationnelle. De plus, notre spécification de τ selon l'équation (4) indique que la part relative des taxes directes payée par le voteur médian varie également selon les taxes indirectes perçues par les gouvernements et les situations de surplus ou de déficit budgétaire dans lesquelles se retrouvent les gouvernements. Enfin, on pourrait également définir un niveau du revenu du voteur médian différant du revenu médian dans l'économie.

À titre d'exemple de cette dernière possibilité, nous pourrions tirer le revenu du voteur médian de la distribution des revenus pour l'ensemble des citoyens exerçant leur droit de vote et construire le rapport des taxes directes associées à ce revenu sur les taxes totales directes perçues dans l'économie. Nous allons ignorer la question des taxes indirectes et du surplus ou du déficit budgétaire et illustrer le rôle joué par la modification concernant le revenu du voteur médian sur l'indice de prix en distribuant le produit national brut de l'économie sur l'ensemble de la population active, même si cette population exclut les rentiers avec droit de vote et inclut les travailleurs ou chômeurs trop jeunes pour voter.[12] En supposant que le voteur médian se trouve parmi la population active et en conservant l'hypothèse de taxes proportionnelles au revenu, alors:

$$\tau = \frac{m^A}{y^A N^A} = \left(\frac{m^A}{y^A}\right) \frac{1}{N^A} \qquad (10)$$

où y^A = P.N.B. $\div$ N^A

N^A = population active.

L'indice A signifie actif(ve).

On peut comparer l'ordre de grandeur de ces hypothèses en utilisant des données compilées sur le rapport du revenu médian sur le revenu moyen et les autres variables impliquées.[13] Ainsi en 1980 par exemple, $\frac{m}{y}$ est égal à y 0,76, la population totale N est de 23,90 millions et la population active A N^A de 11,522 millions. Sous l'équation (9), τ est égal à 0,000000031 et égal à 0,000000065 avec l'équation (10), soit une part relative

des taxes directes totales payée par le voteur médian qui a doublé avec cette dernière spécification.[14] De plus, en utilisant notre définition de y^A, on voit que les revenus médians impliqués sont, pour 1980, de l'ordre de $(\frac{m}{y}) \cdot y = 10\,000$ \$ avec la spécification de la population totale, et de 20 000 \$ $(= (\frac{m}{y}) \cdot y^A)$ pour la spécification de la population active.[15] En principe, nous pourrions calculer deux indices de prix qui seraient associés à ces deux définitions différentes du revenu du voteur médian. En pratique cependant, ce qui nous intéresse, c'est de comparer l'évolution de ces deux indices à partir d'une base commune, comme par exemple, 1971 = 100.

Un des intérêts de la spécification par la population active concerne justement l'évolution de cette variable. Ainsi, une augmentation de la population active va réduire la part relative des taxes directes payée par le voteur médian si la distribution des revenus (telle que mesurée par le rapport du revenu médian sur le revenu moyen) demeure inchangée: l'ensemble de la population continue de consommer les biens publics selon le degré d'encombrement effectif, mais il y a, par ailleurs, plus de citoyens qui contribuent au financement du coût de production de ces biens publics.

Soulignons que cette spécification introduit une dimension démographique dans la théorie de la demande des biens publics par le biais direct du prix relatif des biens publics. Par voie de conséquence, cet élément démographique est également présent dans le calcul d'un indice de prix d'utilité constante incluant les biens publics.

Si nous reprenons l'équation (9) en identifiant le revenu total de l'économie au produit national brut (P.N.B. $= y \cdot N = y^A \cdot N^A$) et en prenant comme mesure du mouvement pur des prix privés l'indice Laspeyres produit par Statistique Canada,[16] nous constatons alors que l'indice des prix à utilité constante de l'équation (7) se réécrit:

$$T^I = 100 \left[\left(1 - \frac{G^o}{\text{P.N.B.}^o}\right) \left(\frac{\text{Laspeyres}^1}{100}\right)^{1-\sigma} + \frac{G^o}{\text{P.N.B.}^o} \left(\frac{(\frac{m}{y})^1 \, q^1 \, N^{1(\alpha^1 - 1)}}{(\frac{m}{y})^o \, q^o \, N^{o(\alpha^o - 1)}} \right)^{1-\sigma} \right]^{\frac{1}{1-\sigma}} \tag{11}$$

puisque, par contrainte budgétaire, $\frac{P^0(1+t^0)X^0}{m^0} = 1 - \frac{T^0}{m^0}$ et que

$$\frac{T^0}{m^0} = \frac{\tau G^0}{m^0} = (\frac{m}{y})^0 \frac{1}{N^0} \frac{G^0}{m^0} = \frac{G^0}{yN^0} = \frac{G^0}{P.N.B.^0}$$

Avec les hypothèses associées à l'équation (10) sur le rôle de la population active, nous obtenons l'indice suivant en utilisant les mêmes substitutions:

$$T_A^I = 100 \left[\left(1 - \frac{G^0}{P.N.B.^0}\right) \left(\frac{\text{Laspeyres}^1}{100}\right)^{1-\sigma} + \frac{G^0}{P.N.B.^0} \left(\frac{(\frac{m}{y})^1 q^1 \frac{N^1 (\alpha^1)}{N_A}}{(\frac{m}{y})^0 q^0 \frac{N^0 (\alpha^0)}{N_A}} \right)^{1-\sigma} \right]^{\frac{1}{1-\sigma}} \quad (12)$$

Pour compléter l'ensemble des données nécessaires à nos simulations, nous avons estimé les coûts unitaires de production des biens publics, q, par l'indice implicite des dépenses publiques courantes en biens et services.[17] Comme cet indice repose sur la convention selon laquelle la productivité du secteur public demeure fixe, nous avons également supposé, comme hypothèse alternative, que q puisse varier au même rythme que la croissance des prix des biens privés, i.e. l'indice Laspeyres de l'I.P.C.

Le tableau 1 résume les résultats de nos simulations. On constate d'abord, à la suite de Lloyd [1975], que les indices basés sur une fonction d'utilité C.E.S. ne varient pas, à toute fin pratique, selon la valeur de l'élasticité de substitution entre biens privés et biens publics. Des exceptions, cependant, concernent les indices spécifiés avec la population active, lorsque le coût unitaire de production des biens publics varie au rythme de l'indice Laspeyres des prix à la consommation (q = L). Pour ces cas, les indices calculés baissent graduellement et de façon continue lorsque l'élasticité de substitution croît pour atteindre

TABLEAU 1. Simulations d'indices de prix à utilité constante incluant les biens produits par les gouvernements (période 1948-1980: quelques années choisies

	Spécification reliée à la population totale								Spécification reliée à la population active							
	$\sigma = 0$				$\sigma = 1,5$				$\sigma = 0$				$\sigma = 1,5$			
	$\alpha = 0$		$\alpha = 1$		$\alpha = 0$		$\alpha = 1$		$\alpha = 0$		$\alpha = 1$		$\alpha = 0$		$\alpha = 1$	
	q = q	q = L	q = q	q = L	q = q	q = L	q = q	q = L	q = q	q = L	q = q	q = L	q = q	q = L	q = q	q = L
1948	57,50	73,51	48,58	58,10	57,44	68,57	47,02	57,98	58,13	74,60	48,96	58,75	58,01	69,14	47,57	58,56
1950	60,81	76,00	52,56	62,22	60,79	72,05	51,04	62,12	62,23	78,38	53,34	63,73	62,13	73,38	52,32	63,45
1955	68,43	80,18	61,49	70,04	68,41	77,77	60,67	69,92	71,48	84,63	63,71	73,29	71,20	80,55	63,40	72,71
1960	75,07	83,32	70,30	77,13	75,06	82,11	69,99	77,00	78,28	87,48	72,96	80,58	78,02	85,08	72,93	79,97
1965	80,42	85,68	77,81	82,61	80,42	85.29	77,68	82,54	83,36	89,16	80,49	85,77	83,23	88,11	80,49	85,36
1970	97,02	98,14	96,57	97,68	97,02	98,13	96,57	97,68	97,64	98,78	97,19	98,31	97,64	98,75	97,19	98,30
1971	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00
1972	105,26	104,43	105,71	104,87	105,25	104,43	105,70	104,87	104,56	103,75	105,01	104,18	104,56	103,74	105,01	104,18
1973	112,82	111,93	113,82	112,91	112,82	111,92	113,81	112,91	110,88	110,03	111,84	110,97	110,84	109,94	111,83	110,93
1974	124,89	123,55	126,73	125,34	124,89	123,53	126,70	125,34	121,69	120,45	123,41	122,11	121,57	120,22	123,38	122,02
1975	142,86	136,72	145,70	139,24	142,69	136,69	145,25	139,24	137,52	131,98	140,08	134,25	137,51	131,54	140,06	134,07
1976	155,31	145,78	159,31	149,15	154,97	145,69	158,45	149,15	148,80	140,30	152,37	143,30	148,80	139,58	152,27	143,01
1977	167,72	156,61	172,91	160,92	167,35	156,46	171,84	160,92	159,71	149,96	164,27	153,74	159,70	148,88	164,17	153,31
1978	179,37	168,25	185,34	173,24	179,24	167,86	184,64	173,22	170,06	160,46	175,21	164,77	169,85	158,58	175,21	163,87
1979	195,17	183,52	202,32	289,53	195,07	183,02	201,54	189,51	183,77	173,93	189,81	179,01	183,34	171,52	189,80	177,87
1980	215,65	201,59	224,50	208,92	215,50	201,05	223,41	208,90	201,79	190,12	204,14	196,20	201,27	187,00	209,13	194,75

Source: imprimés d'ordinateur S1039, S1040, 1013 et 1022.

des différences de taux cumulatifs de croissance entre indices de l'ordre de 3% pour la période de 1971 à 1980.

Conformément aux attentes, le niveau des indices est plus faible après 1971, selon que les biens produits par les gouvernements sont des biens publics purs ($\alpha = 0$) plutôt que des biens privés collectivement financés ($\alpha = 1$). L'ordre de grandeur des différences, en termes de croissance des indices pour la période 1971-1980, tourne autour de 10% en retenant la population active comme spécification, contre environ 8% avec la spécification par la population totale.

Les indices s'avèrent également très sensibles à la spécification sur le coût unitaire de production, q, des biens gouvernementaux. En utilisant l'indice implicite des prix des dépenses publiques en biens et services pour évaluer q, contre l'indice Laspeyres des prix à la consommation ($q = L$), nous obtenons un niveau du premier indice nettement plus élevé après 1971, pour atteindre des différences de taux cumulatifs de croissance de l'ordre de 15% pour la période 1971-1980.

Il faut préciser qu'au cours de cette période, l'I.P.C. s'est accru de 110,6% (indice au niveau 210,6 en 1980) contre 154,4% (indice à 254,4 en 1980) pour l'indice des prix implicites des dépenses publiques. Si le véritable indice de coût (q) devait se trouver quelque part entre ces deux mesures, il et plus vraisemblable qu'il se situe près de l'indice implicite des prix des dépenses publiques. On sait en effet qu'au seul chapitre du travail, par exemple, les employés des gouvernements ont généralement bénéficié d'une indexation à l'inflation selon les mouvements de l'I.P.C. et de rentes de monopoles syndicales qui ont doté ces employés de hausses salariales supérieures au secteur privé au cours des dernières années.[18]

Enfin, également dans des différences de taux de croissance de l'ordre de 15%, on notera que l'indice selon la spécification par la population active montre un taux de croissance inférieur à celui basé sur la spécification se référant à la population totale. La responsabilité de ce résultat incombe au taux de croissance de 33,5% pour la population active pour la période 1971-1980 contre un taux de croissance de 10,8% pour la population totale du Canada. Cette croissance importante de la population active a donc fait décliner le prix

des biens publics pour le voteur médian, dans le cas de la spécification liée à la population active.

Au tableau 2, nous résumons, en termes de taux de croissance du coût de la vie, l'essentiel des résultats obtenus des simulations.

Comme l'indique le tableau 2, la variabilité des résultats de simulation se traduit par une variabilité importante dans les taux d'inflation et dans les calculs du coût de la vie selon les différents indices.

On doit noter, en termes relatifs, l'ordre de grandeur des variabilités que ces simulations suggèrent par rapport aux indices usuels ou par comparaison aux changements associés à d'autres simulations ou expériences suggérées par la théorie des indices de prix. Rappelons, particulièrement sur ce sujet, que le document de référence de Statistique Canada (1982, pp.93, 94) sur les indices de prix montre que, pour l'indice d'ensemble, l'écart qui s'est accumulé sur trois ans, entre l'indice basé sur des biens de 1974 et celui tiré d'une mise à jour du panier de biens de 1978, n'est que de 0,5% en termes relatifs, soit un peu plus de 0,1% par an en moyenne.

Les résultats du tableau 2 montrent que les plus fortes hausses du coût de la vie sont liées à l'indice implicite des dépenses publiques et, ensuite, à l'I.P.C. Les indices incluant les biens publics et les biens privés présentent des taux comparables ou supérieurs à l'I.P.C. dans l'hypothèse de la part relative des taxes directes totales payée par le voteur médian impliquant la population totale; ces taux sont, par ailleurs, inférieurs ou comparables à l'I.P.C. dans l'hypothèse basée sur la population active. Dans ce dernier cas, l'indice le plus comparable en pratique à l'I.P.C., sans évidemment l'être au plan théorique, implique l'indice implicite des prix des dépenses publiques, q, et un niveau d'encombrement dans la consommation des biens gouvernementaux égal à l'unité ($\alpha = 1$).

Les simulations laissent évidemment imprécise la réponse à la question de pointer l'indice de prix jugé le plus près de la réalité économique. Nous croyons cependant que l'approche proposée dans cette étude permet de nous confronter adéquatement à cette question. On peut, en effet, dériver une demande de biens publics, de la part du voteur mé-

... de croissance (en pourcentage) du coût de la vie ou taux d'inflation selon différents indices

Période	Spécification de la population totale $\sigma = 1,5$				Spécification de la population active $\sigma = 1,5$				Indice implicite des dépenses publiques	Indices Laspeyres des prix à la consommation (I.P.C.)
	$\alpha = 0$		$\alpha = 1$		$\alpha = 0$		$\alpha = 1$			
	q = q	q = L	q = q	q = L	q = q	q = L	q = q	q = L	q	I.P.C.
1950-55	2,5	1,6	3,7	2,5	2,9	1,9	4,2	2,9	5,7	2,5
1955-60	1,9	1,1	3,1	2,0	1,9	1,1	3,0	2,0	4,8	2,0
1960-65	1,4	0,8	2,2	1,4	1,3	0,7	2,1	1,3	3,8	1,7
1965-70	4,1	3,0	4,9	3,7	3,5	2,4	4,1	3,0	7,6	4,1
1970-75	9,4	7,8	10,1	8,5	8,2	6,6	8,8	7,3	13,1	8,5
1975-80	10,2	9,4	10,8	10,0	9,2	8,4	9,9	9,0	12,6	10,4

dian, tirée des éléments discutés jusqu'ici et qui nous permet de définir les diverses spécifications et valeurs des paramètres les plus conformes à la réalité statistique.

Si nous maximisons le Lagrangien formé de la fonction d'utilité de l'équation (6) et la contrainte budgétaire résultant d'une substitution des équations (2) et (4) dans (3), nous obtenons l'expression suivante des conditions d'équilibre:

$$\frac{X}{H} = \left[\frac{(1-\delta)}{\delta} \; \frac{p(1+t)}{\tau\, qN^{\alpha}}\right]^{\frac{1}{\eta-1}} \tag{13}$$

En multipliant cette équation par $\frac{p(1+t)}{\tau\, qN^{\alpha}}$ et en réutilisant l'équation (2) ainsi que la contrainte budgétaire telle que spécifiée plus haut, nous obtenons:

$$\frac{(m-\tau G)}{\tau G} = \frac{(1-\delta)}{\delta}^{\frac{1}{\eta-1}} \left[\frac{p(1+t)}{\tau\, qN^{\alpha}}\right]^{\frac{\eta}{\eta-1}} \tag{14}$$

En spécifiant τ selon l'équation (9) ou (10), nous obtenons une expression statistiquement estimable. Par exemple, en retenant la spécification de la population active pour τ et en la substituant dans l'expression (14), nous obtenons, après quelques réaménagements et le recours à la fonction logarithmique:

$$\ln\left(\frac{1-\frac{G}{\text{P.N.B.}}}{\frac{G}{\text{P.N.B.}}}\right) = \beta_0 + \beta_1 \ln\left(\frac{p(1+t)\cdot N^{A}}{(\frac{m}{y})\, q}\right) + \beta_1\,\beta_2 \log N + \epsilon \tag{15}$$

où $\beta_0 = \frac{1}{\eta-1} \ln (\frac{(1-\delta)}{\delta})$.

$\beta_1 = \frac{\eta}{\eta-1}$, de sorte que $1-\beta_1 = \sigma$: l'élasticité du substitution

$\beta_2 = -\alpha$ où α est le paramètre d'encombrement

ϵ = un terme aléatoire aux propriétés usuelles, ajouté de façon ad hoc.

A l'aide de données couvrant la période 1947-1980, nous avons estimé, à titre exploratoire, l'équation (15) de même que l'équation correspondant à la spécification de τ selon la population totale. Plusieurs difficultés ont surgi au niveau économétrique, notamment celle d'obtenir des estimés des paramètres compatibles avec nos spécifications théoriques (par exemple, $0 \leq \alpha \leq 1$). Il faut de plus souligner le problème important de multi-collinéarité entre les variables. Seuls les résultats rapportés au tableau 3, obtenus pour la spécification par la population active et selon des valeurs spécifiques de α, apparaissent convenables.

Marginalement appuyée par la statistique et tenant compte de la valeur du paramètre δ, la régression la plus acceptable est celle comportant une spécification où les biens gouvernementaux sont des biens privés collectivement financés ($\alpha = 1$). Il est à noter d'ailleurs que, pour les études portant sur la demande de biens gouvernementaux, le paramètre d'encombrement dans la consommation des biens publics a toujours été estimé à l'unité.[19] Ces résultats économétriques suggèrent de plus une élasticité de substitution entre biens privés et biens gouvernementaux différente de zéro et près de 1,5.

Par rapport à nos simulations, ces résultats appuieraient l'indice de prix spécifié par la population active, le coût unitaire de production q, et un paramètre d'encombrement égal à l'unité. Or, l'évolution de cet indice a suivi de relativement près l'évolution de l'I.P.C. au cours des dernières années. Toutefois, l'analyse de régression confirme que ce mouvement similaire provient de coïncidences numériques.

TABLEAU 3. Estimation de l'équation (15) par M.C.O. corrigés pour l'autocorrélation des résidus

Hypothèse sur α	Paramètres de la régression		Paramètres structuraux		$\bar{R}^2$	Log de la fonction de vraisemblance
	β_0	β_1	δ	σ		
0	7,87 (6,07)	-0,46 (0,37)	0,99	1,46	0,963	52,73
1	-0,159 (0,38)	-0,616 a (0,36)	0,476	1,61	0,964	53,36

( ): écart-type.
a: significatif à 0,90, test bilatéral.
Source: Imprimés d'ordinateur S1046, S1215.

4. Conclusion

Le modèle du voteur médian a permis d'intégrer, dans une structure théorique, les biens publics dans la construction d'un indice de prix à utilité constante.

De plus, à un niveau agrégé, nous avons démontré qu'il est possible de définir une contrepartie empirique à cet indice. Si l'indice qui en découle sur les bases des données utilisées a produit un mouvement similaire à celui de l'I.P.C. au cours des dernières années, ce résultat apparaît, par ailleurs, comme une pure coïncidence numérique sans autre forme de signification.

Plusieurs lacunes demeurent dans la définition de l'indice. Par exemple, le problème des taxes indirectes pose une question de pertinence de l'indice au plan de la logique économique. Il existe de plus une variabilité importante dans le niveau et l'évolution temporelle de l'indice selon la valeur de certains paramètres et, en particulier, celle de la part de taxes directes totales payée par le voteur médian. Enfin, on ne connaît pas l'opérationnalité de l'indice à niveau désagrégé.

Il demeure cependant que l'approche est théoriquement fondée et pourrait répondre à d'autres préoccupations concernant l'indice de prix et les biens produits par les gouvernements. Par exemple, on pourrait examiner l'effet sur l'indice d'une situation où la production d'un bien passe du secteur privé au secteur public, comme ce fut le cas il y a quelques années pour le secteur de la santé. Avec l'I.P.C. traditionnel, on sort ce bien de l'indice et, si son prix est supérieur à la moyenne, l'indice baisse; dans le cas contraire, l'indice monte et devient dans un cas comme dans l'autre moins représentatif de l'ensemble des prix dans l'économie. Avec notre modèle, nous croyons qu'il serait possible de modifier la fonction de production du bien public H (équation (2)) et de traiter de ce problème comme une question d'économies d'échelles de production selon les secteurs.[20]

Renvois

[1] Gillingham et Greenlees [1983] discutent de l'incorporation des taxes indirectes et directes dans l'indice de prix à la consommation sur la base d'un indice-revenu du coût de la vie en opposition à un indice de dépenses du coût de la vie comme base de référence. Cette procédure mesure le revenu avant impôt nécessaire pour atteindre un niveau donné de satisfaction (dans le sens de l'I.P.C.), compte tenu de l'évolution des taxes. Cette procédure ignore la dimension ''consommation des biens gouvernementaux'' que ces taxes produisent.

[2] Voir Aranson et Ordeshook [1981], pour une critique générale.

[3] Les modèles bureaucratiques et du type logrolling , sont souvent présentés comme des alternatives au modèle du voteur médian. Par ailleurs, ces modèles sont également controversés et se prêtent mal à une discussion sur la théorie des indices de prix. Une exception à signaler, cependant, est l'étude de Bös [1978] qui, à l'aide d'une fonction d'utilité des politiciens, relie les indices du coût de la vie à une politique des prix publics.

[4] Il n'y a pas, à notre connaissance, de discussion sur les taxes indirectes et les surplus ou déficits budgétaires dans la littérature concernant le voteur médian.

[5] $\tau(\ ) = \tau$ pour simplifier la notation.

[6] Voir Lloyd [1975], p.305, équation 14, pour une démonstration précise de cette propriété.

[7] Traduit de l'anglais.

[8] Pour certains paliers de gouvernement, seules les taxes indirectes sont permises.

[9] Le gouvernement pourrait évidemment retrouver l'effet prix en modifiant la part payée par le voteur médian des taxes directes totales, i.e. τ. L'ensemble de ces remarques a évidemment une incidence sur la demande de biens publics et sur l'explication de la part relative des dépenses gouvernementales dans l'économie.

[10] L'expression est d'Aranson et Ordeshook [1981].

[11] Voir Aitchison et Brown [1963].

[12] La population active ne comprend pas les pensionnaires des institutions, les membres des Forces armées, les Indiens vivant dans les réserves, ni les résidents du Yukon et des Territoires du Nord-Ouest. Source: *Revue statistique du Canada*, 11-003F, section 4, tableau 3.

[13] Source: $\frac{m}{y}$: estimés des données de Statistique Canada, catalogue 13-559, y interpolant certaines années durant la période 1951-1973, et par extrapolation, pour 1948-1951, 1974-1980 en régressant les données de S-C sur les coefficients de Gini de Buse [1980] préalablement transformés. N: S-C, 13-001.

[14] On suppose que $\left(\frac{m^A}{y^A}\right) = \left(\frac{m}{y}\right)$.

[15] Source: Y = P.N.B.: S-C, 13-001.

[16] Source: Indice Laspeyres de l'I.P.C.: *Revue statistique du Canada*, S-C, 13-001.

[17] Source: G: S-C, catalogué 13-001
q: 1947-1974, S-C, catalogue 13-531
1974-1980, S-C, catalogue 13-001.

[18] Voir les études de Lacroix et Cousineau [1977] et de Lacroix [1982] sur la question de l'évolution des salaires dans les secteurs public et privé.

[19] Une exception est l'étude de Dudley et Montmarquette [1981] pour l'estimation de la demande de dépenses militaires.

[20] Dans un texte récent, Dudley et Montmarquette [1983] ont spécifié, à l'aide d'une fonction de production de H plus générale que celle définie par l'équation (2), un modèle de demande de biens publics par le voteur médian.

Références

Aitchison, J. et J.A.C. Brown, *The Lognormal Distribution with Special Reference to its Use in Economics*, Cambridge University Press, 1963.

Aranson, P.H. et P.C. Ordeshook, ''Regulation, Redistribution and Public Choice'', *Public Choice*, 37, 1981, pp.69-100.

Bergstrom, T.C. et R.P. Goodman, ''Private Demand for Public Goods'', *American Economic Review*, 63, 1973, pp.280-296.

Borcherding, T.E. et R.T. Deacon, ''The Demand for Services of Non-Federal Governments'', *American Economic Review*, 62, 1972, pp.891-901.

Bös, D. ''Cost of Living Indices and Public Pricing'', *Economica*, 45, 1978, pp.69-79.

Bowen, H.R., ''The Interpretation of Voting in the Allocation of Economic Resources'', *Quarterly Journal of Economics*, 58, 1943, pp.27-48.

Buse, A., ''The Cyclical Behavior of the Size Distribution of Income in Canada'', Discussion Paper No. 80-13, University of Alberta, 1980.

Dudley, L. et C. Montmarquette, ''The Effects of Non-clearing Labor Markets on the Demand for Public Spending'', manuscrit non publié, janvier 1983.

__________ ''The Demand for Military Expenditures: An International Comparison'', *Public Choice*, 37, 1981, pp.5-31.

Gillingham, R. et J.S. Greenlees, ''The Incorporation of Direct Taxes into a Consumer Price Index'', *La Mesure du Niveau des Prix: Actes du Colloque tenu sous l'égide de Statistique Canada,* 1983.

Goldberger, A.S., ''Functional Form and Utility: A Review of Consumer Demand Theory'', University of Wisconsin: Systems Formulation, Methodology and Policy Workshop Paper 6703, octobre 1967.

Hotelling H., ''Stability in Competition'', *Economic Journal*, 39, 1929, pp.41-57.

Kessel, R.A., ''The Measurement and Economics Implications of the Inclusion of Indirect Taxes in the Consumer Price Index'', dans *The Price Statistics of the Federal Government*, N.B.E.R., 1961, chapitre 12.

Klein, L.R. et R. Rubin, ''A Constant-Utility Index of the Cost of Living'', *The Review of Economic Studies*, XV, 1948-49, pp.84-87.

Lacroix, R., *Les conséquences des disparités de salaires entre secteurs privé et public au Québec*, dossier # 16, Conseil du patronat du Québec, mars 1982.

________ et J.M. Cousineau, *La détermination des salaires dans le monde des grandes conventions collectives: une analyse des secteurs privé et public*, Conseil économique du Canada, Ottawa, 1977.

Lloyd, P.J., ''Substitution Effects and Biases in Non True Price Indices'', *American Economic Review*, 65 (3), juin 1975, pp.301-313.

Statistique Canada, Document de référence de *L'indice des prix à la consommation*, catalogue 62-553, hors série, mai 1982, 151 pages.

PRICES, PROXIES AND PRODUCTIVITY: AN HISTORICAL ANALYSIS OF HOSPITAL AND MEDICAL CARE IN CANADA

Morris L. Barer
Health Services Research & Development
Robert G. Evans
Department of Economics
University of British Columbia

SUMMARY

This paper assembles data from several sources to construct price indexes for medical nd for hospital services in Canada from 1935 to 1980. The CPI components for the years rior to the public insurance programs are linked with an aggregate of medical fee and enefit indexes compiled for each province by the Department of National Health and Velfare, and with data from the hospital statistical reports to Statistics Canada. These ndexes are then combined with estimates of total expenditures in each category, as well s other data on personnel and system capacity, to derive estimates of the growth of "real" utput in each sector.

The medical fee index is, however, based on list fees, and must be adjusted for signifiant increases in collections ratios over this period. Direct information on these ratios was ot available, but indirect estimates were derived from data on insurance coverage and hysician supply. While list fees rose 3.20% per year, on average, from 1935 to 1979, our stimates of growth in fees collected range from 4.18% to 5.01%. Overall CPI growth as 3.95%. "Real" services per physician and per capita, based on list fees, rose 3.20% nd 4.65%, with greatest apparent growth in the period of introduction of public insurance; djustment for collections reduces these to 1.41% – 2.22% and 2.84% – 3.66% respectively.

For hospital care, an index of input prices is constructed from 1960 to 1980, based primari- on an index of labour costs. This input price index rose 7.78% per year on average, mpared with the CPI at 5.35%, primarily due to rapid wage gains in the hospital sector .36% p.a. relative to 7.43% p.a. for the industrial composite, and 6.97% for services). his index is then linked in 1960 to the CPI hospital services component, extending it back

to 1935. The resulting index of hospital prices grows 6.94% per year, compared with the CPI of 4.08%, a growth in relative prices for hospital care of 2.75% per year for 45 years!

Hospital expenditures per admission and per patient-day are then divided by this price index, yielding a residual "service volume index" representing intensity of servicing. The growth in hospital expenditures of 13.10% p.a., 1935-80, can thus be factored into 1.78% population increase, 1.46% patient-days per capita, and 9.52% expenditures per day. The latter, in turn, is composed of 4.08% CPI, 2.75% hospital specific inflation, and 2.41% increase in servicing intensity per patient-day.

One cannot, however, interpret this as "quality" improvement in the absence of information on either the efficacy of care or the informed choices of consumers. The paper concludes with a discussion of the conceptual problems of interpreting health care prices.

RÉSUMÉ

La composante "Santé et soins personnels" de l'IPC est trompeuse, en ce sens qu'elle exclut les dépenses en soins hospitaliers et médicaux – les principale composantes des dépenses en services de santé. Les programmes d'assurance publique ont transféré ces dépenses des budgets privés aux budgets publics à la fin des années 50 (hôpitaux) et des années 60 (médecine), et l'IPC a été corrigé en conséquence. En fait, cependant, l'inclusion de ces formes de dépenses dans l'IPC est douteuse même en l'absence d'assurance publique. L'importance très réduite du choix du consommateur dans les dépenses de santé, et la prédominance très large d'assurances-santé privées non basées sur des données actuarielles (et souvent obligatoires) ont pour effet de modifier le statut des dépenses de santé d'allocations budgétaires à une forme plutôt capricieuse d'impôt, même dans le cas d'un système privé.

Néanmoins, le comportement des prix dans le secteur des soins de santé est d'une importance considérable pour la politique des soins de santé et également pour la politique économique générale. L'interprétation des "explosions de coûts" ou du "sous-financement" dans les soins de santé dépend évidemment du fait que les mouvements des dépenses représentent principalement des effets de prix ou de quantité. On voudrait savoir quelles

sont les tendances sous-jacentes des volumes réels des services. De même, le "coût d'opportunité" représenté par un niveau donné de dépenses de santé, relativement à l'IPC, dépend nettement de la mesure dans laquelle il représente les coûts réels en ressources, ou simplement une redistribution de fonds entre les fournisseurs, les utilisateurs, les contribuables ou les cotisants. Et les effets des assurances, privées ou publiques, sur l'utilisaion des soins de santé, par opposition aux prix, posent un problème important auquel on ne peut pas répondre d'une façon globale en l'absence de statistiques de prix décrivant les périodes de changement. En dépit de la grande quantité de données recueillies sur les composantes des dépenses, et sur d'autres aspects de notre système de soins hospitaliers t médicaux, nous manquons de données sur les prix qui soient cohérentes, continues, et omparables d'une période et d'une province à l'autre.

Ce texte rassemble des données de diverses sources pour établir des indices de prix pour es services médicaux et hospitaliers au Canada entre 1935 et 1980. Les composantes de 'IPC pour les années qui ont précédé les programmes d'assurance publique sont reliées un agrégat d'indices d'honoraires et de bénéfices médicaux préparés pour chaque proince par le ministère de la Santé nationale et du Bien-être, et à des données sur les rapports statistiques des hôpitaux de Statistique Canada. Ces indices sont alors combinés avec es estimations des dépenses totales dans chaque catégorie, et aussi avec d'autres données ır le personnel et la capacité du système, pour calculer des estimations de la croissance la production "réelle" de chaque secteur.

L'indice des honoraires médicaux est, cependant, basé sur des listes d'honoraires, et doit re corrigé pour des augmentations importantes des taux de perception au cours de cette ériode. Les renseignements spécifiques sur ces taux n'étaient pas disponibles, mais des timations indirectes ont été calculées à partir des données sur le champ d'application s assurances et le nombre de médecins. Alors que les listes d'honoraires ont monté de 20% par année, en moyenne, entre 1935 et 1979, notre estimation de l'augmentation s honoraires perçus s'étale de 4.18% à 5.01%. La croissance générale de l'IPC a été 3.95%. Les services "réels" par médecins et par habitant, basés sur les listes d'honoraires, ıt monté de 3.20% et de 4.65%, et la croissance apparente la plus importante se situe ıns la période de mise en vigueur de l'assurance publique; la correction pour les taux perception réduit ceux-ci à 1.41% – 2.22% et 2.84% – 3.66% respectivement.

Pour les soins hospitaliers, un indice des prix des entrées a été construit pour la période allant de 1960 à 1980, indice basé principalement sur un indice des coûts de la main-d'oeuvre. Cet indice des prix des entrées est monté en moyenne de 7.78% par année, comparativement au taux de 5.35% de l'IPC, principalement à cause des hausses rapides des salaires dans le secteur hospitalier (9.36% par année comparativement à 7.43% pour l'indice industriel composite et à 6.97% pour les services). Cet indice est alors relié en 1960 à la composante des services hospitaliers de l'IPC, et on le ramène rétrospectivement jusqu'à 1935. L'indice des prix hospitaliers qui en résulte augmente de 6.94% par année, comparativement à une augmentation de 4.08% de l'IPC, une augmentation des prix relatifs des soins hospitaliers de 2.75% par année pendant ***45 ans!***

Les dépenses des hôpitaux par admission et par jour/malade sont alors divisées par cet indice de prix, produisant un indice résiduel du "volume des services", qui est une mesure d'intensité des services. La croissance des dépenses des hôpitaux de 13.10% par année entre 1935 et 1980, peut ainsi être ventilée en une augmentation de la population de 1.78%, une augmentation de 1.46% des jours/malades par habitant, et une augmentation de 9.52% des dépenses par jour. Ce dernier chiffre est lui même composé de 4.08% de l'IPC, 2.75% de l'inflation spécifique aux hôpitaux, et d'une augmentation de l'intensité des services de 2.41% par jour/malade. On ne peut cependant pas interpréter ceci comme une amélioration de la "qualité", en l'absence d'information sur l'efficacité des soins ou la possibilité de choix réel des consommateurs.

Une caractéristique frappante des séries des prix hospitaliers et médicaux est l'importance de la rupture des tendances antérieures qui se produit au cours des années 70, et principalement les cinq dernières années. Les listes d'honoraires des médecins n'ont pas suivi la montée de l'inflation pour la première fois depuis le début des années 50, et elles retardent de 4.7% par année dans les années 1971-75 et de 1.14% entre 75 et 79. Au cours des périodes antérieures de diminution des listes d'honoraires "réels" entre 1935 et 1954, la hausse des taux de perception semble avoir gardé les honoraires perçus en hausse au rythme de l'IPC ou légèrement plus vite, alors que la facturation réelle par médecin a également augmenté. Mais après 1971, les taux de perception sont stables. Au début des années 70, la facturation réelle par médecin a monté de 3.08%, ce qui compensait presque la diminution des honoraires des listes. Mais à la fin des années 70, ce processus semble également avoir cessé. Pour la première fois en une génération, la facturation des médecins semble

avoir été déterminée par les listes d'honoraires, ce qui expliquerait les difficultés accrues de la négociation de ces honoraires.

Du côté des hôpitaux également, la fin des années 70 marque une rupture des tendances antérieures. L'indice des salaires hospitaliers, qui dépasse les salaires généraux entre les années 1960 et 1975, prend un léger retard par rapport à eux en 1975 et 1980, ce qui suggère que la longue période de croissance du revenu relatif pour les travailleurs hospitaliers peut avoir pris fin. De même, l'indice des prix hospitaliers "réels", les prix des entrées des hôpitaux par rapport à l'IPC, se stabilisent après 1975 pour la première fois depuis la Seconde Guerre mondiale.

Les taux d'utilisation par habitant, stables entre 1970 et 1975, ont fait un bond entre 1975 et 1977, mais semblent s'être stabilisés de nouveau, et l'indice d'intensité des services par jour/malade, qui a monté de presque 4.0% en moyenne entre 1952 et 1971, est demeuré au-dessus de 2% au cours des années 70, et presque toute l'augmentation s'est concentrée en 1976.

Les indices de prix exposés ici fournissent ainsi des informations sur le contexte des conflits politiques de plus en plus intenses sur les ressources de soins de santé au cours des années 80.

. Introduction

In 1981 Canadians spent nearly 8% of their total gross national income, over $26 billion, n health care services. This percentage, while large, is relatively low in international terms; number of other OECD countries currently spend in the neighbourhood of 10%. Cana-an experience was markedly unusual during the 1970s, our spending share fluctuating etween 7% and 7.5% from 1970 to 1979 while most of the rest of the world raced ahead. the previous two decades the Canadian experience was much like everywhere else; health ending rose steadily from about 4% of GNP just after the war to over 7% in 1970. ırallels with the U.S. are particularly striking; the relation of health care expenditure GNP is virtually identical in both countries from 1950 to 1971, and diverges sharply ereafter.

As in the case of most statistics, what these numbers reveal is interesting, but what they conceal is critical. In particular their normative interpretation is currently the focus of heated debate and advocacy, some of which spills over into the public domain, but much of which is behind closed doors. The percentage by itself indicates that something important has happened, that the public health insurance programs, which were fully in place by the early 1970s, have permitted global expenditure control. The common U.S. argument that public insurance, and government funding generally, must lead to cost escalation, is false as a general proposition. But these numbers do **not** tell us whether such control is good or bad policy, and much of the current debate is over precisely that.

Providers of health care, in particular, are arguing that the Canadian system is ''underfunded'', that insufficient resources are being devoted to health service production, and that the public insurance program is the source of the difficulty. Those responsible for payment, by contrast, ''point with pride'' to the success of the public program (at least up until 1980) in containing costs which in other countries such as the U.S. are totally out of control. The significance of the debate is enormous; the future of the country's health care system could be at stake. Yet obviously the percentage figure **by itself** tells us nothing at all about whether the system is relatively efficient, or starved for funds.

To address this issue, we need two further classes of information. One of these will be obvious to economists, but the second may be less familiar. First, and the focus of this paper, is price data. An expenditure series obviously factors into price and quantity components, and the focus of health policy is (usually) quantities. The ''underfunding'' argument is ostensibly a claim that inadequate quantities of services are being provided, though it may in fact be a cover for a claim for higher provider prices/incomes. It is then critically important that one be able to separate out price and quantity effects over time, and determine to what extent expenditure escalation or control represent movements in relative prices as against shifts in real volumes of services provided/utilized. A number of subsidiary issues in health care policy turn on the same point; one must be able to factor expenditures by sub-component of health care and by geographic region, as well as over time, into price and quantity components in order to draw inferences about the effects of past policies and the scope of future ones.

Secondarily, and equally important, expenditure and resource policy must be soundly based on efficacy information. "Underfunding" implies that **needed** services are being foregone, which means that the health or well-being of consumers-patients is suffering as a result. It is thus a technical statement, about the existence of efficacious interventions not now being utilized because resources are insufficient. This grounding, common alike to critics and advocates of the present funding system, is quite distinct from the usual economist's criterion of consumer sovereignty. The test is not that there are interventions which consumer-patients are willing to undergo, either at present zero money prices or at prices equal to their marginal resource costs, whatever those might be, but whether such nterventions would lead to health benefits. This alternative evaluative standard, peculiar o health care, is not the explicit focus of this paper, but inevitably underlies the interpretation of both price and quantity data.

The development of price data in the health care sector does not begin to keep pace vith its economic and social significance. In particular the Consumer Price Index (CPI) eflects almost none of the movements of health care prices. The health care sub-index as a weight of only 1.415 in the overall CPI, and covers a small group of health care oods and services which are themselves a small and unrepresentative part of total health are spending. Table 1 makes the point clearly; the CPI health care component is almost xclusively an index of the cost of dental care and drugs. It covers items which make up nly about 17% of total health care spending. Indeed the share of these items in gross ational expenditure, 1.26%, is not far off their weight of 1.41 in the overall CPI, despite ny differences in concept and measurement which may exist between categories with similar bels.

The almost total exclusion of health care expenditures from the CPI is ostensibly based n the fact that the vast bulk of such expenditure passes through public budgets rather an private pockets. Data from family expenditure surveys show that in 1978 average family penditure on health care was only $391.20, or under 2% of family spending, but even is is too high. It includes health insurance premiums, both for the public programs in ose provinces (Ontario, Alberta, British Columbia) which still retain that form of finance, d private insurance premiums. But the public plan premiums, being effectively comulsory and actuarially unrelated to expected utilization, are more properly treated, as the ational Accounts do, as a form of taxation. Out-of-pocket spending net of insurance

premiums is only $249.20 per family, but this excludes both private insurance premium and out-of-pocket payments subsequently reimbursed by private plans. Insofar as sucl private plans are primarily employee groups, however, the individual employee may hav little or no choice as to participation. The only difference between public and private grou insurance is the form of public subsidy. Public plans, in provinces using premiums, mak up the difference between premium income and plan outlays from general revenue. Privat premiums must cover total outlays, plus a load factor, but are tax-subsidized in that, unlik public plan premiums, employer-paid private premiums are not taxable in the hands o the employee.

TABLE 1. Health Care Expenditures in Canada, 1978, by Major Component, and in Relation to CPI Components

	$billion	%	% of CPI Component
TOTAL HEALTH CARE EXPENDITURE	$16.18	100.0	100.0
Institutional Care	8.61	53.2	---
Hospitals	7.34	45.4	---
Professional Services	3.68	22.7	52.5
Medical	2.54	15.7	---
Dental	0.92	5.7	52.5
Drugs and Appliances	1.83	11.3	47.5
Prescribed Drugs	0.84	5.2	20.6
Non Prescribed Drugs	0.76	4.7	9.2
Eyeglasses	0.16	1.0	15.6
Other appliances	0.06	0.4	2.1
Other Expenditures	2.07	12.8	---

The grounds for exclusion of health care costs from the CPI are, however, broader tha those which were in fact invoked by the Dominion Bureau of Statistics. Component dele tion historically followed the spread of public plans; thus cities in Saskatchewan were drop ped from the Medical Care index in 1962 and their weights reassigned to other cities. Mos other cities were dropped when they entered the federal program in 1968, and by 197 the medical care component of the CPI was based on data from Quebec and New Brunswick By 1971 the component had vanished. The same procedure had been followed 10 year before with respect to hospital care, the component referring to fewer and fewer cities ove time until by 1961 it was terminated entirely.

No one could quarrel with the dropping of these components, once hospital and medical spending had been transferred almost entirely to public budgets. There are, however, two important issues which do arise. First, given the peculiar nature of health care as a commodity and of its typical financing mechanisms even in the absence of public programs, should it ever have been in the index, and if so in what form? Secondly, after health spending largely dropped out of the CPI, where did it go? It disappeared into the government ector of the GNE deflator. But the individual identity of the health care sector was swallowed up in that mass, and as a result we now have seriously incomplete and inadequate price lata with which to address important issues of health policy.

The two problems come together, as we shall suggest at the end of the paper. There nay be an important class of goods and services – education being another example – which re collectively provided yet individually consumed. They are not purchased in the private narketplace, but at the same time are not pure public goods like defence or foreign affairs. These collectively-provided services do display marked price variations over time and egions, and these price changes have important welfare effects and policy implications. et our systems of price statistics do not incorporate them in any meaningful way.

The inappropriateness of the CPI as a framework for health care prices in the presence f private insurance has already been touched on. The conceptual problems this creates e most evident in the case of a province like British Columbia, where by 1968 most of ie population already had medical care insurance of a form essentially similar to the public ans. The public program simply took over the files and the personnel of the private inirers. Yet the CPI reacted as if **all** of medical care had suddenly shifted from private public payment. The point is that in the presence of extensive private health insurance, *fortiori* if that insurance is community-related and tax-subsidized, health care expenture has already left the private marketplace. The changes which take place in its costs e very little more under the control of individuals or families than are their taxes. Private ental insurance makes this point as well. Once a group of employees enters a contract, dividuals must accept and pay for coverage as a condition of employment. They can odify their obligations only through political action, by attempting to persuade a marity of the group to try to change the contract.

Lying behind this problem, however, is the peculiar nature of health care itself, however funded. Prior to the public plans, the CPI treated health care as equivalent to any other item on which the consumer family might choose to spend its resources. This implicitly assumed that health care was a ‘‘good’’ like food, clothing, or shelter. More recently, however, students of health care have come to realize that health **care** as opposed to health itself, is not a ‘‘good’’ in the normal sense, a conclusion which enables one to understand the very peculiar institutions which surround and regulate its production and consumption in every developed society.

The boundaries around the health care sector are rather fuzzy; food, clothing and shelter obviously also contribute to health, and probably in a more important way than health care. What distinguishes health care goods and services is that their anticipated effect on the user’s health status is the principal, often the only, reason for use. Food may affect our health (for good or ill), but we do not eat, at the margin, solely or primarily with health effects in mind. By contrast, we go to the physician, and *a fortiori* the dentist, only because we believe that our health status will be improved as a result.

Indeed, health care *per se* has negative consequences, as the dentist example makes clear. In general, health care use has two effects on our well-being (quite apart from its cost). Insofar as care contributes positively to health, and health is of value, health care thus contributes indirectly to well-being. But its direct effects, holding health status constant, are in most cases negative. Diagnostic and therapeutic interventions are uncomfortable and frightening. They are undergone, not consumed, in the expectation of health benefit.

Formally, the consumer’s utility function is:

$$U(X_1,\ldots,X_n,HS(X_j,\ldots X_m,HC),HC)$$

where utility depends on consumption of various goods X_i, but also on health status HS and consumption of health care HC. A subset m of all other goods n (m may equal n) also has effects on health status.

In this form, $dU/dX_i > 0$ -- goods are goods -- and $dU/dHS > 0$ also -- being healthy s better than being ill. But $\partial U/\partial HC \lesseqgtr 0$, and in general < 0 -- health care hurts. It iffects health, however, so

$$dU/dHC = (\partial U/\partial HC) + (dU/dHS)(\partial HS/\partial HC) > 0.$$

'he positive indirect effects of care outweigh its negative direct effects, or it will not be onsumed -- **if** patient-consumers are fully informed. The subset of X_i in m may have ositive or negative effects on HS, so long as their total effects on U are positive.

It is readily apparent from this formulation that the relationship between health status nd health care is a production relationship nested in a utility function. The usual assump-on that the consumer has privileged access to knowledge about the structure of the utility ınction and the marginal rates of substitution among its various arguments cannot be pplied to the expression dU/dHC, because the critical component $\partial HS/\partial HC$ is a technical ıdgment about the efficacy of care. Hence, the extensive regulatory structure such that nly licensed, self-regulating physicians may provide medical services and admit to hospitals. he physician's gatekeeper role is only justifiable on the ground that his/her superior nowledge as to the structure of the HS function leads to better choices of HC levels and atterns than would the decisions of consumers transacting at arms length from unregulated, rofit-maximizing providers.

This approach has certain parallels with the characteristics model of demand introduc-l by Lancaster, but it also has important differences. In particular, HS is not a ıaracteristic embodied in health care; it is a state specific to an individual, and the effect health care on any particular individual depends on what that individual's initial state . Care which is expected, on balance, to improve the health of one person may easily wer that of another in a different state. Antibiotics for victims of bacterial and of viral fections would be one common example, toxic drugs or radical surgery for cancer would a more extreme case.

This extended view of the underlying utility function implies that one cannot readily fer utility from consumption behaviour. The role of the provider inserts an additional rm in the linkage, breaking the theoretical connection between the utility function and

observed consumption behaviour. It would not be inconsistent, for example, for increases in utilization to follow **increases** in relative prices, as utilization represents not a point on a stable demand curve, but a resolution of the forces bearing on provider and patient alike. The theoretical consequences of this situation have been developed in detail elsewhere (Evans and Wolfson, 1980; Evans, 1982); they raise serious questions about the appropriateness of including medical care in the consumer price index even if it were all paid out-of-pocket. Again the analogy is with taxes, which are also paid out-of-pocket, yet are not assumed to constitute an item of utility-enhancing expenditure. They are, rather, "regrettables" -- the cost of the social overhead, like defence, police and fire insurance; and they arise in connection with unfortunate, utility-reducing events over which consumers have little or no control.[1]

To illustrate the problem another way, think of two families of equal size, confronting the same prices. One spends twice as much as the other on food, and both have identical resources left over after paying for food. If we have no other information, we would guess that the family which spent more on food is better off (happier) though it apparently has different tastes as well. But its members can purchase the same consumption bundle as the other family, and more, if they choose.

Now for food, substitute automobile gasoline. Again, without additional information we would guess that the high gas-consuming family was better off. If additional information were to indicate that this family in fact uses that extra gas to commute 60 miles per day to and from work, we might be inclined to question the family's ''better off'' status. But since we are assuming that all prices (including housing) faced by the two families are equivalent, we must conclude that the gas-guzzling family does derive additional utility from the process of commuting, or from benefits made available by commuting, since it has made employment and housing choices from which derive the need to commute.

Finally for food, substitute hospital or medical care. Our judgment as to relative well-being is now reversed, because a reasonable inference would be that the high-spending family was sicker. Although they can buy the same bundle of non-medical goods as the low-spenders, they will still be worse off because of the direct disutility effects of illness, which are not fully compensated by care. Of course if we knew that both families had had the same illness experience, we might expect the high-spenders to be better off, but lacking

that information, we would have to take high health care spending as an inverse index of well-being. A family that spends a lot on health care is no more to be envied than a community that spends a lot on police or defence. The spending is both without **direct** utility (and in the case of health care at least, generates direct disutility) and is also a signal for the presence of other, direct sources of disutility -- illness, crime, foreign conflict.

The same problem intrudes into our interpretation of the quantity series which we derive from expenditure and price data in the following sections of this paper. It is customary o refer to increases in real service volume per physician as increases in productivity, im-licitly holding constant the quantity of other inputs per physician and average physician-ours of work. Thus total expenditures on physician services, deflated by an appropriate rice index and divided by the physician stock (measured in full-time equivalents), will f all are correctly measured, yield an index of ''physician productivity''. But in this field, ecause of the nesting of a production relation in the utility function, patient well-being annot be inferred from use. The true linkage is:

Resources -------> Services ---> Health Status -----> Utility

he measurement of physician productivity referred to above applies only to the first two erms in this series. It does not follow, but is critically important for policy purposes, that xtra services translate into improved health status. If they do not, then the resources are asted. The highly efficient production of laboratory tests which convey no diagnostic formation and do not affect therapy is not output.

More questionable, and more difficult to classify, is the efficacious service which leaves tility unchanged or lowered -- the procedure is more uncomfortable or otherwise unplea-nt than its expected benefits justify. Again, however, it seems plausible to argue that creased output of such services is not really improved productivity in a broader sense.

Of course, fully informed consumer-patients, if they existed, would not use services which ere inefficacious or almost so, since for such services $dU/dHC < 0$. In practice, however, source use in this field is guided by physicians who are likely to underestimate, or even sregard on professional grounds, the $\partial U/\partial HC$ term. More important, they suffer from stematic and documented upward biases in their estimates of $\partial HS/\partial HC$, such that utiliza-n of services of unproven effect or known lack of effect is widely observed. The user wever, **believes** that $\partial HS/\partial HC > 0$.

The same issue arises in hospital care. After adjusting for input price change, we find in Section III below that real resources per hospital admission or per patient-day have increased substantially over time. If we call this reduced productivity, we are assuming that each admission or patient-day represents the same amount of service over time. In the physician case we implicitly assumed that a physician represented an intertemporally constant amount of capacity, and we measured observed changes in service flow per unit of capacity. In the hospital case we would be assuming a constant service definition and observing changes in resource flows per unit of service. This leads to an apparent steady **increase** in physician productivity, and **decrease** in hospital productivity, both of which may indeed be so. But one should keep in mind that hospital facilities and personnel are non-priced inputs to the physician's practice. We are also assuming that ''other inputs'' to the physician's practice have remained constant per physician, which may be plausible for directly employed inputs (overheads have been roughly stable at least since the 1950s). But the growth of hospital spending may represent an expansion of hospital-supplied inputs to medical practice. We may be observing diversion of resources to increase ''productivity'' where it will yield larger incomes for physicians.

Disregarding this possibility, there is still a serious issue in the assumption that a hospital admission or patient-day is the ''same'' unit of output from 1935 to 1980. Doubtful as that proposition may be, however, it is equally fallacious to assume that its effectiveness increases in proportion to the resources used -- an implicit assumption of constant productivity of resources. This is done, for example, if resources per day or case are used as a ''quality'' index. What one of course needs to know is, once again, the relation of services to health status. If hospital episodes do in fact become more efficacious as their intensity of resource use increases, then clearly output in some broader sense is rising, and the real input per episode series underestimates productivity increase, or rather, overstates its decrease. But without more efficacy information we cannot tell. And there is no reason in theory or practical experience to give the institution the benefit of the doubt. (There are some issues concerning the amenity dimensions of institutional care, whereby one might move directly from resources to utility without the two intervening steps. But in the context of the actual hospital utilization process these serve rather to muddy than to clarify the central question of productivity.)

We return to the issues of interpreting implied service volume indexes after describing their development for medical and hospital services respectively.

II. The Medical Services Sector

The index of medical care prices developed in this section of the paper is based on data from Statistics Canada/DBS and the federal Department of National Health and Welfare. These are the only sources of which we are aware that have collected and prepared national time-series data on physicians' fees.

DBS, as noted above, included physicians' services as one of the items in the "market basket" on which the CPI was based. Physicians were therefore surveyed in the cities used in constructing the CPI, and reported the prices they charged for an office call, a home visit, a confinement and an appendectomy. The latter two are inclusive charges -- for all associated medical care. These four items were then combined into a single Doctors' Fees Index for each city, and then aggregated to the national level.

The most complete published source for this index appears to be in *Prices and Price Indexes* (Vol. 47 No. 1, January 1969, in Table II on pp. xv and xvi). This reports the health care component of the CPI, together with its sub-indexes of physicians' and dentists' fees, optical care, pre-paid medical care, and pharmaceuticals, from 1949 to 1968, on a base of 1961 = 100. The physicians' fees components for 1969 and 1970 are given in the December 1969 and December 1970 issues of *PPI*, on pages 50 and 51 respectively. These data, adjusted to 1971 = 100, make up the 1949 to 1963 values of the Physicians' Fees Index, MDFI, reported in column 1 of Table 3.

The earlier years, back to 1935, were provided as unpublished data by Statistics Canada, Consumer Prices Division, as a "Doctor" index based on 1949 = 100 which overlapped with the published series (except for the difference in base value) in the years 1949 to 1968. These were linked to column 1 of Table 3 by rebasing to 1971.

The geographic coverage of this index is not as wide as the CPI generally, and varies over time. In particular, as noted above, provinces with universal public medical care programs are dropped from the index, and their weights reassigned to other provinces. Thus

Saskatchewan, or rather Regina and Saskatoon, are dropped from the index midway through 1963, Vancouver midway through 1968, and most other regional cities by the end of 1968. In its last year, 1970, the index refers only to cities in Quebec and New Brunswick.

Going back in time, however, the coverage of the index also becomes less complete. While prices of the same four items were surveyed, at least as far back as 1949, particular cities may not have been included throughout. Information received at different times from Statistics Canada on this point is inconsistent and has not been reconciled.

The other source of national data on physicians' fees is the Department of National Health and Welfare, which assembles the fee schedules issued by provincial medical associations and the benefit schedules of the provincial reimbursement agencies. The compilation of fee schedules goes back to December of 1963. From these are developed measures of the overall percentage change in fees resulting from each fee schedule modification. This provides a base from which one can develop an index of fees for each individual province; the Department also prepares comparisons of fee schedule levels across provinces at particular points in time. The Policy, Planning and Information Branch of DNHW has developed from this a national "Medical Care Plan Fee-Benefit Index" annualized from 1971 to 1981, using underlying monthly data.

The "Fee-Benefit Index" has to wrestle with the problem of periodic discrepancies between the schedules promulgated by provincial medical associations, and those approved for reimbursement purposes by provincial paying agencies. In the late 1970s, several medical associations issued recommended fee schedules well above the provincially approved rates, to indicate what they believed their members ought to be paid and (in some cases) to encourage physicians to opt out of the provincial program and to charge patients the higher fees. The DNHW index is based on changes in the public reimbursement schedules, and thus does not respond to changes in the relative prices of opted-out or extra-billing physicians, or to the spread of these practices. It will thus be slightly low in the late 1970s and early 1980s, particularly in Alberta and Ontario. The overall significance of direct billing in total medical expenditures is alleged to be small, and we had no hard data available on which to base upward adjustments to the provincial reimbursement schedules.

A more significant problem arises in the pre-1971 period, in which provinces were setting up public medical care insurance programs and adopting benefit schedules for the first time. These were generally taken over from the medical association schedules, which were previously used by the private not-for-profit insurers. But in some cases governments adopted benefit schedules which were discounted from the medical association schedule, in an attempt to adjust for the fact that historically, collections ratios were below 100%. Medical association fee schedules were a ''list price'' from which physicians were expected to discount explicitly or implicitly for patients less able to pay; thus average prices received ran below list prices by an amount which would vary by region, specialty and individual practitioner.

The DNHW fee index does not, however, attempt any specific adjustment to the fee indexes in each province when the public programs were established. It treats this point as a clean break. Our MDFI series, on the other hand, passes directly over the break by carrying forward the previous medical association schedule as if it were unchanged. This amounts to assuming that any adjustment for a change in the collections ratio by a percentage markdown from the private fee schedule to the public benefit schedule was in fact an accurate adjustment, on average, so that actual fees collected, as opposed to listed, were not affected by the creation of the public plans.

We are quite sure, of course, that this assumption is invalid. Its justification, however, is that collections ratios, and hence the relation between actual and list fees, appear to have been changing over the whole 1935-71 period. Private insurance, both service-benefit and indemnity, grew rapidly from 1945 to 1968, and in total probably had substantially more effect on collections ratios than the public plans. The move to public programs was merely part of a long process of adjustment which in some provinces (those where extra-billing is still practised) is not over yet. Accordingly we chose to smooth over the period of entry to Medicare, and to try to adjust for changing ratios of actual to list fees by a more comprehensive process described below.

In this context it should be noted that changes in collections ratios are not the only source of variation in the relationship of actual to list fees. Fee schedules were not legally binding; individual practitioners could charge above or below the recommended rate. Such schedules were introduced in the late 1950s and early 1960s, so there was probably a steady

drift in each province during the pre-Medicare years towards increasing conformity to the schedule. Anecdotal evidence from the more recent experience of dentists suggests that new practitioners are more likely to adopt a professional guide, while established practitioners are slower to modify their independent fee-setting behaviour. Profession-wide fee co-ordination takes place over time. But such drift will be upwards, as the professionally promulgated schedule is likely to be at or above average list prices of practitioners.

Furthermore, billing patterns may change. ''A la carte'' billing and procedural reclassification can lead to significant increases in actual fees received for a given set of activities, with no change in officially recorded fees. The classic case is Quebec, where from 1971 to 1975 the public benefit schedule was unchanged while payments per practitioner rose 4% per year. Such ''strategic billing'' occurs in all provinces to a greater or lesser degree, and appears to have been more prominent in the early days of fee schedule development and negotiation. As reimbursers have become more experienced in identifying the items of greatest potential for such cost increases, one would expect the negotiation process to lead to schedules which provide fewer opportunities.

A further problem with the DNHW data is that prior to 1971 they do not include Quebec. Since there **was** no single provincial fee schedule for Quebec, the exclusion is understandable -- it also justifies DNHW's reporting of their national index only from 1971 onwards.

Offsetting these problems, the DNHW index has much wider coverage than the CPI component. In each province it now includes some hundreds of fee schedule items, and considerable care and attention has been devoted to ensuring comparability of items across provinces. Furthermore, the weights with which these items are combined are developed from actual utilization and cost data, on a current basis, implying both that such indexes will be more sensitive to shifts in practice patterns and that they are more appropriate for deflation purposes. And issues of the relationship between such indexes and a theoretical ''constant utility'' index do not arise, for reasons discussed above in Section I.

In constructing the MDFI series in Table 3, we took the province-specific percentage changes in fee schedules reported by DNHW back to the base of December 1963, and thus generated an index for each province relative to 1971 as a base of 100.0. The index for each year took account of the month of introduction of any fee schedule change; thus

a province which had no change during 1971 or 1969, but a 10% change on April 1st of 1970 would have an index value of 1971 = 100, 1969 = 90.9, and 1970 = 97.7. If a fee change occurred in 1971, the average fee level effective throughout 1971 was set equal to 100.0.

These provincial indexes were then aggregated into a nine-province "national" index from 1964 to 1970, using the provincial utilization weights reported by DNHW for the period closest to 1971, with the Quebec weight assigned equiproportionately to the other nine. The result was an index which overlapped the CPI component from 1964 to 1970. The results are as shown in Table 2.

The two indexes are remarkably close, despite their substantial differences in origin and mode of construction, suggesting that physicians surveyed by the compilers of the CPI were reporting list fees, and on average following the medical association fee guides quite closely. One can therefore link the CPI component and the DNHW index with some confidence that they are measuring the same things and share the same biases.

The actual linkage was done by comparing the CPI component in column 1 of Table 2 with the DNHW index in column 2 carried out to two more digits (not reported) and finding that their ratio varied between 1.3742 and 1.4132, with no apparent trend, and an average of 1.4033. But the 1969 value of 1.3742 is a clear outlier, and the 1969-70 period is suspect in any case, because by that time the CPI component referred only to Quebec and New Brunswick, and the DNHW index excludes Quebec. An attempt was made to modify the DNHW-based index by using Montreal CPI data from 1963 to 1970, but the results were not encouraging. The set of six physicians surveyed in Montreal over that period appeared to display atypical pricing behaviour. Accordingly the average was taken from 1964 to 1968, yielding a value of 1.4073, and this figure was used to adjust (deflate) the CPI component for 1963 and all prior years.

TABLE 2. Alternative Physicians' Fees Indexes, Canada, 1964-70

Year	CPI Component 1961 = 100	DNHW Index 1971 = 100	CPI 1967 = 100	DNHW 1967 = 100
1964	107.0	76.1	87.4	87.9
1965	110.1	78.7	90.0	90.9
1966	112.7	80.1	92.1	92.5
1967	122.4	86.6	100.0	100.0
1968	127.8	90.6	104.4	104.6
1969	132.0	96.1	107.8	111.0
1970	138.2	97.8	112.9	112.9
% change 1964-70	29.2	28.5		

In summary, then, the MDFI index in Table 3 is patched together from the DNHW 10 provinces' public insurance fee benefit index from 1971 to 1981, a nine-province fee schedule index aggregated by the authors from DNHW reports for 1964 to 1971, and the CPI physicians' fees component, based on 1961 = 100 and then divided by 1.4073, from 1935 to 1963.

The first point to notice about this series is that it behaves very differently from the CPI. In Table 3, the CPI is similarly based at 100.0 in 1971, and the ratio of MDFI to CPI is shown along with the annual percentage changes in MDFI, CPI and MDFI/CPI. It is clear that very long swings have occurred in the relationship between the list prices, at least, of physicians' services, and the price level of consumer goods generally. In relative terms, physicians' list prices were at their peak at the beginning of the period, in 1935, and from the trend, may have been even higher in the depths of the depression. After 1935, the ratio of list fees to CPI fell in almost every year, except for a brief up-tick at the end of the war, to reach a trough in 1951, 27.1% below its 1935 value. In spending power terms, physicians' fees dropped to less than three-quarters of their initial value, or about 2% per year over the 16-year period. Most of the drop was during the recovery from the depression (1935-37) and the post-war inflation (1946-48), suggesting that "sticky" physicians' fees were outpaced in a period of rising prices generally.

TABLE 3. Physicians' Fees Index

Year	MDFI	CPI	MDFI/CPI	Annual % Change		
				MDFI	CPI	MDFI/CPI
35	37.7	34.8	1.08			
36	37.6	35.5	1.06	−0.26	1.94	−2.16
37	37.7	36.6	1.03	0.26	3.17	−2.82
38	37.5	37.0	1.02	−0.39	1.02	−1.40
39	37.4	36.7	1.02	−0.39	−0.81	0.43
40	37.4	38.2	0.98	0	4.09	−3.93
41	38.2	40.4	0.94	2.08	5.89	−3.60
42	38.4	42.3	0.91	0.51	4.64	−3.94
43	38.6	43.0	0.90	0.63	1.77	−1.12
44	39.8	43.3	0.92	3.15	0.52	2.62
45	40.4	43.5	0.93	1.35	0.52	0.82
46	41.1	45.0	0.91	1.93	3.45	−1.47
47	42.8	49.2	0.87	4.02	9.33	−4.86
48	46.3	56.3	0.82	8.19	14.48	−5.49
49	48.7	58.0	0.84	5.15	3.06	2.03
50	49.1	59.7	0.82	0.88	2.84	−1.91
51	52.1	66.0	0.79	6.08	10.55	−4.05
52	55.9	67.6	0.83	7.37	2.50	4.75
53	57.4	67.0	0.86	2.67	−0.89	3.59
54	58.7	67.4	0.87	2.23	0.56	1.66
55	59.5	67.5	0.88	1.45	0.22	1.23
56	61.2	68.5	0.89	2.74	1.44	1.28
57	63.6	70.7	0.90	3.95	3.17	0.75
58	67.1	72.6	0.92	5.48	2.65	2.75
59	69.0	73.4	0.94	2.86	1.14	1.70
60	69.9	74.3	0.94	1.34	1.23	0.11
61	71.1	75.0	0.95	1.63	0.91	0.71
62	73.2	75.9	0.96	3.00	1.20	1.78
63	74.5	77.2	0.97	1.84	1.78	0.06
64	76.1	78.6	0.97	2.12	1.75	0.37
65	78.7	80.5	0.98	3.36	2.48	0.86
66	80.1	83.5	0.96	1.81	3.72	−1.85
67	86.6	86.5	1.00	8.13	3.59	4.38
68	90.6	90.0	1.01	4.57	4.07	0.48
69	96.1	94.1	1.02	6.06	4.50	1.49
70	97.8	97.2	1.01	1.86	3.35	−1.44
71	100.0	100.0	1.00	2.20	2.85	−0.64
72	101.4	104.8	0.97	1.40	4.80	−3.24
73	102.3	112.7	0.91	0.89	7.58	−6.22
74	107.4	125.0	0.86	4.99	10.90	−5.34

TABLE 3. **Physicians' Fees Index** – Concluded

Year	MDFI	CPI	MDFI/CPI	Annual % Change MDFI	CPI	MDFI/CPI
75	114.2	138.5	0.82	6.33	10.79	−4.03
76	121.8	148.9	0.82	6.65	7.47	−0.76
77	132.0	160.8	0.82	8.37	8.01	0.34
78	140.2	175.2	0.80	6.21	8.95	−2.51
79	150.6	191.2	0.79	7.42	9.16	−1.59

MDFI: Physicians' Fees Index Linking CPI Medical Care Component and Health and Welfare Canada Fee-Benefit Schedule Data.

Source: See Text.

From 1951 to 1969, however, the trend was reversed. List fees rose, in relative or real purchasing power terms, in almost every year, to reach a peak 29.3% above their 1951 low. During this 18-year period of generally rapid growth in real incomes, and more-or-less (at least in retrospect) general price stability, the list fees of physicians rose 1.44% per year in real terms.

But at the end of the 1960s, the trend was reversed yet again. A combination of direct negotiation of list fees with provincial bargainers, plus rapid and incompletely anticipated inflation, plus federal wage and price controls for part of the period, brought list fees down sharply in real terms between 1971 and 1974, a drop of about 5% per year for three years. Since then the fall has been less precipitous but still steady, at about 1.4% per year, to a level in 1981 23.8% below the 1969 peak. In real terms, physicians' list fees relative to the CPI were in 1981 just below their previous trough of 1951. It remains to be seen whether the bargaining of 1981 and 1982 will have reversed this trend a third time.

We have emphasized throughout, however, that the MDFI series is strictly a measure of list fees, not an index of fees actually charged, much less collected. Moreover we know that the ratio of list to actual has been falling over time, with rising incomes and the spread of public and private insurance, until by the late 1970s and early 1980s the national average is probably slightly below unity. Accordingly the actual levels of physicians' fees collected, which one should use to deflate expenditure or physician income series to measure aggregate utilization or service output per physician, have been rising more rapidly than MDFI. To indicate the order of magnitude of this effect, we present in Table 4 the results of a deflation of estimates of total expenditure on (fee-practice) physicians' services, assuming that MDFI were a valid price index.

The first column, MDEXP, is drawn from periodic reports by the Department of National Health and Welfare, published and unpublished, of national health care expenditure and its components. This set of reports goes back to 1960, but the less inclusive personal health care expenditures series goes back to 1951 and includes physicians' services on the same basis as NHE. From there back to 1935, data were available from unpublished material prepared by DNHW for inclusion in a revised version of Urquhart and Buckley's *Historical Statistics of Canada*. There is a difficulty, however, in that data from different reports for the same concept and period generally do not quite match. We have therefore relied upon whatever appeared to be the most recent report. For dates prior to 1970 this was the unpublished series from DNHW, but readers of the published reports will note that discrepancies are small.

Table 4 also includes a series of data on the total physician stock -- ACP or active civilian physicians, not fee-practice physicians, to compare with apparent "real service output" estimates. This series is derived from various issues of the *Canada Health Manpower Inventory*, from 1969 on. From 1963 to 1969 it is available in annual reports of DNHW on the Canadian health care system. Judek, in his study for the Hall Commission (Table 2-2, p. 25), reports active civilian physicians from 1951 to 1961, based on various CMA surveys;

the 1961 to 1963 bridge is based on unpublished DNHW data and looks a bit shaky. Prior to 1951, data are from Urquhart and Buckley's Historical Statistics of Canada for 1941, 1943, and 1947-1950 (series B108), with geometric interpolation by the authors. The 1931 census is used as a base point for interpolating from 1935 to 1941.

Table 4 then derives two series, MDXR and WLMD, which are respectively ''real'' output of medical care services and workload per MD. The former is MDEXP/MDFI, and the latter is MDXR/ACP. It should be emphasized that MDXR is not ''really'' real; rather it is our calculation of the total volumes of medical services production and utilization implied by an uncritical acceptance of MDFI as a true index of physicians' fees. WLMD introduces an additional complication. Active civilian physicians include a number, perhaps as many as a third, of physicians who are working on salary, full- or part-time, in medical practice, hospitals, administration, research, or teaching. MDEXP is based only on the activities of predominantly fee-for-service practitioners. Thus the levels of WLMD should not be expected to equal real gross billings per fee-for-service practitioner, much less per full-time practitioner. It is the changes in WLMD which are significant.

As a similar ''conditional'', Table 4 presents the Canadian population from 1935 to 1979, and MXRPC, ''real'' medical services **per capita** over time, on the assumption that MDFI is valid.

TABLE 4. Medical Services Per Capita and Per Physician

Year	MDEXP(000)	MDXR(000)	ACP	WLMD(000)	POP(000)	MXRPC
35	43,800	116,260.6	10,723	10.84	10,845	10.72
36	44,500	118,423.5	10,907	10.86	10,950	10.81
37	50,100	132,983.0	11,094	11.99	11,045	12.04
38	54,800	146,024.3	11,284	12.94	11,152	13.09
39	57,100	152,747.3	11,477	13.31	11,267	13.56
40	62,800	167,995.3	11,673	14.39	11,381	14.76
41	66,700	174,785.8	11,873	14.72	11,507	15.19
42	68,900	179,632.9	11,746	15.29	11,654	15.41
43	68,600	177,724.8	11,620	15.29	11,795	15.07
44	66,000	165,762.5	12,011	13.80	11,946	13.88
45	76,200	188,838.2	12,414	15.21	12,072	15.64
46	86,700	210,795.0	12,831	16.43	12,292	17.15
47	91,000	212,691.4	13,263	16.04	12,551	16.95
48	101,400	219,053.8	13,373	16.38	12,823	17.08
49	117,000	240,369.8	13,873	17.33	13,447	17.88
50	135,000	274,943.5	14,099	19.50	13,712	20.05
51	153,000	293,745.0	14,325	20.51	14,009	20.97
52	168,000	300,413.1	15,135	19.85	14,459	20.78
53	176,600	307,585.1	15,829	19.43	14,845	20.72
54	188,600	321,327.6	16,431	19.56	15,287	21.02
55	206,500	346,784.9	17,221	20.14	15,698	22.09
56	240,100	392,442.1	17,871	21.96	16,081	24.40
57	271,795	427,370.8	18,523	23.07	16,610	25.73
58	301,337	449,227.0	19,096	23.52	17,080	26.30
59	325,689	472,033.6	19,800	23.84	17,483	27.00
60	355,014	507,735.9	20,517	24.75	17,870	28.41
61	388,305	546,462.0	21,290	25.67	18,238	29.96
62	406,075	554,823.1	23,248	23.87	18,583	29.86
63	453,395	608,257.3	24,082	25.26	18,931	32.13
64	495,657	651,152.1	24,847	26.21	19,290	33.76
65	545,056	692,785.6	25,481	27.19	19,644	35.27

TABLE 4. Medical Services Per Capita and Per Physician – Concluded

Year	MDEXP(000)	MDXR(000)	ACP	WLMD(000)	POP(000)	MXRPC
66	605,200	755,555.6	26,528	28.48	20,015	37.75
67	686,189	792,274.6	27,544	28.76	20,378	38.88
68	788,089	870,124.3	28,209	30.85	20,701	42.03
69	901,435	938,427.8	29,659	31.64	21,001	44.68
70	1,040,739	1,063,628.3	31,166	34.13	21,297	49.94
71	1,250,413	1,250,413.0	32,942	37.96	21,569	57.97
72	1,386,199	1,367,060.2	34,508	39.62	21,848	62.57
73	1,483,421	1,450,069.4	35,923	40.37	22,125	65.54
74	1,659,651	1,545,298.9	37,297	41.43	22,479	68.74
75	1,914,089	1,676,084.9	39,104	42.86	22,727	73.75
76	2,103,199	1,726,764.4	40,130	43.03	23,025	75.00
77	2,309,011	1,749,250.8	41,398	42.25	23,280	75.14
78	2,535,608	1,808,564.9	42,238	42.82	23,493	76.98
79	2,822,600	1,874,236.4	43,192	43.39	23,701	79.08

MDEXP : Total Expenditure on Physicians' Services.
MDXR : "Real" Volume of Physicians' Services: MDEXP·100/MDFI.
ACP : Number of Active Civilian Physicians.
WLMD : Workload per Physician: MDXR/ACP.
POP : Population.
MXRPC : "Real" Volume of Physicians' Services Per Capita: MDXR/POP.

The results are striking. Over the period 1935 to 1979, the real volume of medical services utilized appeared to rise by 6.5% **per year**, for a total increase of just over sixteenfold. Over this same period, MDFI rose by 3.2% per year. Total expenditures on medical care services rose by 10% per year, on average, for all of this 44-year period; and if we believe the MDFI index, two-thirds of this was real volume increase. Meanwhile, population rose 118.5% from 1935 to 1979, or 1.8% per year, so real per capita service utilization rose 4.6% per year -- for 44 years!

While this would be a rather startling expansion, one cannot rule it out *a priori*. But when we turn to the WLMD series, we find apparent annual growth of real output per MD from 1935 to 1979 of 3.2%. This would imply an increase in productivity substantially greater than the general rate of technical progress of the whole economy for a full generation. Furthermore, by coincidence the average annual rate of increase in the physician stock was also 3.2% per year, indicating that over this period the growth of medical services output drew in equal parts on expansion of the physician stock, and increasing productivity of that stock. The implications for manpower planning would be extreme -- if the data were valid.

Analysis of sub-periods shows, as expected, a dramatic gain in ''productivity'' in the late 1960s. From 1966 to 1971, WLMD rises 5.9% per year, indicating the severe downward bias of MDFI over the period of introduction of the public Medicare plans. But this period is by no means the sole source of difficulties. If we break out the sub-periods, 1935-49 and 1949-61, the annual growth of WLMD averages 3.4% and 3.3% over these periods as well. Even the period 1971-76 shows a very respectable 2.5% annual gain; only after 1976 does WLMD go flat.

Looking at the annual changes in WLMD, it seems clear that the dramatic growth rates of the later 1930s were a catching-up of actual to list fees. As the economy recovered, physicians' collections ratios apparently did likewise. The sharp post-war pattern suggests that while consumer prices generally were outrunning listed fees, physicians' actual billing patterns were changing too. This could reflect the growth of private health insurance. Fast increases in productivity in the late 1950s coincide with the introduction of public hospital insurance, which could have resulted in one or more of an increase in real productivity as

more co-operative inputs became available, a change in billing practices, or an increase in collections as patients were relieved of hospital bills. Finally, the great jump in "productivity" in the late 1960s clearly coincides with the introduction of the public medical care programs, and MDFI obviously fails to capture the associated increases in fees received per service.

To derive a more satisfactory index of fees actually collected, we articulate a relatively simple model of the relationship between list and actual fees, and use the model to estimate actual fees. If we define F as the "real" index of physicians' fees actually collected, the deflator as it exists in the mind of God, and further assume that our total expenditure series is accurate, then

$$X = FQ = F^{*}\hat{Q} \qquad (1)$$

where X is our series MDEXP from published and unpublished sources, F^{*} is our index MDFI, $\hat{Q}$ is our biased estimate of the real volume of service output, MDXR, and Q is the "true" volume index of physicians' services.

Since a varying proportion of the Canadian population over this period had insurance coverage against the cost of physicians' services, we may write

$$F = iF^{*} + (1\text{-}i)kF^{*} \qquad (2)$$

Here i designates the proportion of the population with comprehensive medical care insurance which by assumption pays full list fees; the uninsured pay some proportion k of list fees. The actual fees collected are the average of F^{*} and kF^{*} weighted by the respective proportions of the population. The introduction of the public Medicare program in each province brings i to 1.0, and F to F^{*}.

From (1) we know that

$$\hat{Q} = QF/F^{*}$$

and from (2),

$$F/F^* = i + (1-i)k$$

so substituting $u = (1-i)$ and $\ell = (1-k)$, we get

$$\hat{Q} = Q(1-u\ell) \tag{3}$$

But u is known; we have unpublished data from DNHW reports giving the percentage of the population with comprehensive, or with any, coverage against medical costs. Ideally we might like to know the proportion of expenditure covered, but that is unavailable. $\hat{Q}$ is simply MDXR as calculated. Q and ℓ are unknown.

We can, however, test some patterns which Q and ℓ might follow. In particular, Q will be dependent on the supply of physicians available. As an approximation, we may write

$$\ln Q = \alpha_0 + \alpha_1 \ln M + \alpha_2 \ln B + \alpha_3 t \tag{4a}$$

where M is equal to the ACP series reported above, B is the stock of acute care hospital beds, and t is the time trend. Ideally one would like to use a series such as FTE physicians in fee-for-service practice, and some more comprehensive measure(s) of hospital capital stock and personnel, as well as information on personnel and capital in physicians' private practices, and physician-hours of work. Although (4a) is a very crude proxy for a medical services production function, with the very limited data available over this period it is at least a place to start.

Alternatively, we can construct this function as:

$$\ln(Q/M) = \beta_0 + \beta_1 \ln(B/M) + \beta_2 t \tag{4b}$$

which focuses attention on the volume of output per physician as a function of beds available per physician and time trend.

We can similarly postulate that ℓ, that proportion of medical bills which is left unpaid by people without insurance coverage, is also a function of other observable variables characterizing the users of health care and the factors affecting their payment behaviour. As only three possibilities from a large class, we specify:

$$\ln\ell = \gamma_0 + \gamma_1 \ln Y + \gamma_2 \ln RF + \gamma_3 t \tag{5a}$$

$$\ell = \delta_0 + \delta_1 Y + \delta_2 RF + \delta_3 t \tag{5b}$$

and

$$\ln\ell = \zeta_0 + \zeta_1 \ln Y + \zeta_2 \ln RF + \zeta_3 \ln UH + \zeta_4 t \tag{5c}$$

In these formulations, Y is real per capita disposable income and is expected to exert a negative effect on ℓ -- with rising real incomes, the proportion of medical bills uncollected should fall. RF or relative fees represents the ratio of physicians' list fees to the general price level as represented by the CPI and so might cause collections ratios to fall as it rises. The time trend, t, is as usual a measure of our ignorance or lack of imagination. The final variable, UH, is a measure of the proportion of hospital expenditures paid out-of-pocket, covered by neither private nor public hospital insurance, nor grants. As it falls, the overall financial burden of illness on the ill falls, and one might expect physicians' collections ratios to rise.

Using these components, we can fit by non-linear least squares the equation

$$\ln\hat{Q} = \ln Q + \ln(1-u\ell) \tag{6}$$

where the unknowns Q and ℓ are substituted for by 4a or 4b, and 5a, b, or c respectively. Tables 5 and 6 report results derived from five of these estimations. In these tables, MDXR and MDFI represent $\hat{Q}$ and F^*, while MDXRA, B, etc. and MDFA, B, etc. are the outcomes of attempts to estimate "true" Q and F respectively using, successively, specifications 4a and 5a, 4b and 5a, 4b and 5b, 4b and 5c, and 4a and 5c. In the fifth trial, represented by MDXRE and MDFE, the estimation process failed to converge with both t and UH in the equation, so t was dropped.

It should not be imagined that we believe these exhaust all the possible specifications of (6), in terms of either variables or functional form. In particular, some better measure of hospital capital stock and co-operant resources for the Q function would obviously be worth testing, as would any other time-series data on inputs to physician service production. And the ℓ function cries out for better specification. It may be that ℓ rises as u falls; the residual uninsured being those with least resources and hence most likely to default. We regard the regressions reported here as a first test of an approach to the problem which could be much refined in future.

The procedure followed was to fit equation (6), then calculate the implicit estimate of Q. Annual average growth rates of these different Q estimates are then shown, over selected time periods (related to major institutional changes in public health insurance) in Table 5. Growth rates for the population and the physician stock are shown for the same periods. In Table 6, the implied annual average rates of increase of true or collected physician prices are calculated by applying (1), $F = (F^{*}\hat{Q}/Q)$, and these are shown in comparison with rates of change of the CPI.

The coefficients of the equations themselves are not shown. The series used are highly collinear, and the coefficients are rather unstable, and in any case there was no obvious appropriate test for significance in the non-linear structure. We did, however, calculate for each equation the squared deviations between $\hat{Q}$ observed and the value of $(1\text{-}u\ell)Q$ implied by that equation, summed across all years. It is not clear what absolute standard of goodness of fit should be applied, but it seems to us that in judging among different specifications, the relative ability to track $\hat{Q}$ should count in a specification's favour. The sums of squared deviations are of course large, but if we divide them by 45 (for 1935 to 1979) and take the square root of the result we get an annual average deviation of "predicted" from observed Q of:

specification A: 177

specification B: 221

specification C: 205

specification D: 195

specification E: 271

which may give some idea of the relative reliability of the estimates.

TABLE 5. Average Annual Rates of Growth (%), in Volume of Physicians' Services (Alternative Estimates), Numbers of Active Civilian Physicians, and Population 1935-79 and Sub-periods

Year	MDXR	MDXRA	MDXRB	MDXRC	MDXRD	MDXRE	ACP	POP
1935-79	6.52	4.74	5.52	4.99	4.68	4.79	3.22	1.79
1935-45	4.97	1.54	4.05	3.21	2.75	1.57	1.48	1.08
1945-54	6.08	4.83	5.94	4.93	4.37	4.93	3.16	2.66
1954-60	7.92	5.94	6.43	5.55	5.06	6.05	3.77	2.64
1960-71	8.54	6.83	6.28	6.18	6.13	6.86	4.40	1.73
1960-65	6.41	6.90	6.32	6.22	6.16	6.93	4.43	1.91
1965-71	10.34	6.78	6.24	6.16	6.11	6.81	4.37	1.57
1971-79	5.19	4.95	5.21	5.21	5.22	4.94	3.44	1.19
1971-75	7.60	6.62	5.82	6.16	6.35	6.59	4.38	1.32
1975-79	2.83	3.31	4.60	4.27	4.09	3.32	2.52	1.05

TABLE 6. Average Annual Rates of Growth (%), in Prices of Physicians' Services (Alternative Estimates) and Consumer Price Index, 1935-79 and Sub-periods

Year	MDFI	MDFA	MDFB	MDFC	MDFD	MDFE	CPI
1935-79	3.20	4.95	4.18	4.70	5.01	4.90	3.95
1935-45	0.69	4.09	1.58	2.41	2.87	4.06	2.26
1945-54	4.25	5.49	4.39	5.39	5.96	5.39	4.99
1954-60	2.96	4.88	4.40	5.27	5.76	4.78	1.64
1960-71	3.31	4.96	5.51	5.61	5.66	4.93	2.79
1960-65	2.39	1.92	2.48	2.57	2.63	1.89	1.62
1965-71	4.91	8.41	8.96	9.04	9.09	8.38	3.68
1971-79	5.25	5.49	5.23	5.23	5.22	5.50	8.44
1971-75	3.38	4.33	5.12	4.78	4.60	4.36	8.48
1975-79	7.16	6.66	5.35	5.68	5.86	6.65	8.40

Looking first at Table 5, the MDXR or $\hat{Q}$ series suggests a dramatic 6.52% annual averag increase in the output of physicians' services over the 44- year period, or a sixteenfold tota increase. This implies a 3.20% annual increase per physician, and 4.65% per capita. Thes are very large increases in productivity and real output to be sustained over such a lon period of time; if true they imply a quadrupling of output per physician and an almos eightfold increase in **per capita** use.

The attempts to measure "true" Q by specifications A to E, while they each yield dif ferent results, nevertheless cluster on much more conservative estimates of productivit and utilization increase. The two "best fits" in terms of squared deviations of predicte from observed $\hat{Q}$, are A and D, and these yield the lowest annual average increases in Q but except for specification B, all cluster within one-third of a percentage point. Specifica tion A implies that real service output per physician grew at an average rate of 1.47% pe year, while the outlier B raises this to 2.23%; these look more plausible in terms of th usual estimates of economy-wide productivity growth rates. It is of course true that w are measuring only output per physician, not total factor productivity, but we do not hav data on other inputs hired by physicians. Most of the other input growth, however, too place in the hospital, and this is partially proxied by the hospital bed series. Further wor is obviously indicated.

The sub-period analysis is also enlightening. The alternative estimates seem to be pa ticularly divergent in the late depression and war period, where the data are most doubtf and the general economic disruptions probably most severe. All indicate, however, signi cant overestimates in the output growth implied by official list price series over this peric -- physicians' collections ratios were obviously improving. The apparent productivity i crease of 3.5% per annum implied by MDXR drops to a high of 2.5% and a low of abo zero in this period.

In the early post-war period, with private insurance and public hospital constructic spreading, the growth in physician supply and in population speeded up. Once agai though, all specifications except B suggest that collections ratios continued to improve su that "true" productivity increase was in the range of 1.17% to 1.72%, not the nearly 3 implied by list prices.

The 1954-60 period saw the introduction of public hospital insurance, and another ap-arent jump in physician productivity as well as numbers -- apparent output increases of 00% per year. Most Q estimates suggest that physician productivity did increase -- as ould be expected given the expansion of the hospital sector -- but not so fast.

The early sixties are a time of further acceleration in the growth of physician stock, but etty good agreement between the $\hat{Q}$ series and the various Q estimates. Medical service itput and **per capita** use were growing fast, but physician productivity growth was in e 2% per year range, reasonable since hospital employment and facilities per bed also ew rapidly in this period. But in the 1965-71 period, as public medical insurance was troduced in most provinces, the Q and $\hat{Q}$ series diverge sharply. Apparent average an-ial "productivity" increases of 5.72% per physician per year over six years are rather ubtful; the "true" Q estimates of 1.67% to 2.34% are far more plausible.

Finally in the 1970s, insurance coverage, and presumably collections ratios, stabilize, d the rate of growth of physician stock is cut sharply. The Q and $\hat{Q}$ series should move gether, and generally do, remaining discrepancies indicating problems of equation fit. itput per physician continues to rise at 1.46% to 1.70% per year on average depending on equation chosen. But the first and second half of the 1970s show quite different pat-ns. From 1971 to 1975 the various estimates diverge somewhat, but all show output per ysician rising in the neighbourhood of 2% per year. But in the later period, only the iant B specification touches 2%, and list prices -- which should be accurate over this iod, imply virtually no productivity gain at all after 1975. The 1971-75 period suggests haps strategic billing in the early Medicare years, being choked off later by increasingly erienced payment agencies. Perhaps by the late 1970s the loopholes of improved col-ions, strategic billing, and increases in cooperant inputs from hospitals have all been sed, and physician incomes are finally linked to fee increases (unless hours of work are eased). If this story is true, it would explain the increases in physician militancy in the 1970's and early 1980's.

n the other hand, it should be recalled that ACP is an all-civilian physicians' series, d as a proxy for full-time-equivalent physicians in fee-for-service practice because of vailability over the whole time period. Shifts in the ratio of ACPs to fee-practice-sicians could lead to errors in the reported "productivity" changes.

Table 6 presents the other side of the coin from Table 5, the "true" price increases im plied by the Q series estimated from specifications A through E. What it suggests is tha true fees, or fees collected, outran list fees over the whole 1935-79 period, by from 1% to 2% per year, with the better specifications, A and D, yielding the larger differentials While list fees lagged behind the CPI by nearly three-quarters of a percentage point pe year, fees collected ran ahead by at least one-quarter percent and more likely about 1% per year. Using specification A, prices collected for physicians' services rose 52.4% relativ to the CPI from 1935 to 1979.

Again the sub-periods are crucial. From 1935 to 1945 list fees were almost flat, but co lections rose enough that fees received probably matched or substantially outpaced th CPI. In the post-war inflation, again list fees lagged somewhat, but fees received ran aheac

From 1954 to 1971, however, list fees outran the CPI by a substantial margin, but wer themselves far out-distanced by actual collections. Again using specification A, from 195 to 1960 fees collected rose 3.19% per year faster than the CPI; from 1960 to 1965, the slowed to 0.30% per year faster (slower than the rise in list fees -- suggesting a proble with the equation fit); and from 1965 to 1971, 4.56% per year above the general price leve

From 1971 on the trend has of course reversed. The 1971-75 period was particularly stri ing, with list fees running 4.7% per year behind the CPI. This is the sharpest sustain movement in the historical record, matching the late sixties jump (though sustained f only four years, not six). From 1975 on, list fees have moved up sharply, but still ran behi the CPI by over 1% per year from 1975 to 1979.

What this exercise suggests is that dramatic movements in physicians' fees, relative t the general price level, are not just a feature of the period of introduction of Medicar Sustained periods of rapid increase, as well as flat spots, show up elsewhere in the pr Medicare record, and would be badly distorted by a reliance on list fee data alone. Eve more misleading would be an assumption that physicians' fees collected moved with th general CPI. We are well aware that Tables 5 and 6 are only a first crude cut at studyi the behaviour of medical services prices and outputs in the pre-Medicare period, but w believe that the value of the effort has been demonstrated.

III. The Hospital Sector

By any commonly accepted definition, the hospital sector dominates expenditure on health care. In 1979, institutional care accounted for 54.5% of total health care expenditures, and hospital care for 42.7% (Health and Welfare Canada, 1982a). But definitions, or in- clusions and exclusions, take on some importance in historical examinations of the sector. Data become (not surprisingly) increasingly spotty as one attempts to reconstruct the pre- insurance financial life of Canadian hospitals, and the straightforward (even if exasperating) problem of missing data is compounded by myriad less explicit definitional changes, shifts n category inclusions, and inconsistencies across sources. Even today, Statistics Canada livides hospitals into six categories: public general, public allied special, public mental, public tuberculosis, proprietary, and federal, while Health and Welfare Canada categorizes hospitals as general and allied special, mental, tuberculosis, and federal, and then adds homes for special care (essentially nursing homes) as making up the institutional sector. But of course public general and allied special excludes a small private sector included in he broader Health and Welfare categories. To compound the problem, many of Statistics Canada's historical series are based on 'reporting', not all operating, hospitals. While this eems in most cases to be a minor problem, the fact remains that the percentage and mix of reporting hospitals does change over time.

It is not our intent in this section to attempt the penultimate reconciliation of all historical Canadian hospital data series. Rather, this discussion serves as background if not justifica- ion for our choice of historical series in the analyses which follow. These analyses (except where otherwise noted) are based on the Statistics Canada category, public general and llied special (PGAS) hospitals which in 1980-81 accounted for 84.6% of all operating hospitals, and 96% of total rated bed capacity of operating hospitals (Statistics Canada, *Hospital Statistics, Preliminary Annual Report, 1980-81*). Operating expenses for the 1,049 PGAS hospitals in fiscal year 1979-80 were in the area of $7.3 billion (*ibid*), or about 2.8% of GNP.

Price Indexes

A. Pre-universal Hospital Insurance

During the pre-universal insurance era, hospital rates were maintained as a (largely unpublished) component of the health and personal care component of the CPI. The Statistics Canada series *Prices and Price Indexes* entered the game rather late, commencing regular publication of a hospital rates index with calendar year 1955 (although selected rates do appear in the occasional earlier publication). Since there was no remaining justification for the inclusion of hospital care in a consumer price index (as discussed in Section I above) after the last province entered the public hospital insurance program, this series was terminated after 1960. Statistics Canada has, however, made available an unpublished "hospital rates" series for the period 1935-60 which overlaps the published figures for 1955-60. The entire series appears as column 1 in Table 7.

There are, of course, other pseudo-price series which have at times enjoyed relatively uncritical deployment as price indicators. The attraction of per diems ($/day) or costs per case/separation/admission is that these series can be extended back to the early 1930s and are readily available (some niggling intertemporal inconsistencies aside) in the post-1960 period. Unfortunately as developed below, that is often all that makes them attractive as price series.

Table 7 compares the "hospital rate", "per diem", and "cost per admission" series over the early period leading up to the universal hospital insurance program. The per diems shown there are drawn from Statistics Canada publication 83-203, *Hospital Statistics, Vol. II -- Financial*, 1957, which reports only 1932 and 1944-53 data. We applied the average annual rate of increase 1932-44 to compute a figure for 1935. Comparable 1954-60 figures are taken from Lefebvre (1976, Table 9). Cost per admission was computed as ($ per day)(total patient-days)/(total admissions). Total patient-days, adults and children was computed for the period 1932-52 from series reported in Urquhart and Buckley.[2] For 1953-60 data are from *Hospital Statistics*, 1971: (average daily number of in-patients)x365(or 366). Data on admissions are from Urquhart and Buckley for the 1932-60 period.[3]

TABLE 7. **Hospital Rates and Other "Price" Series, Canadian Public General and Allied Special Hospitals, 1932-60**

Year	Hospital Rate Index	CPI	Real Hospital Rate Index	$ Per Day	$ Per Admission
32				2.73	48.30
35	51.80	59.95	86.41	3.09	54.16
36	52.40	61.11	85.75		
37	53.20	63.05	84.38		
38	53.70	63.70	84.31		
39	54.60	63.18	86.42		
40	54.60	65.76	83.03		
41	55.40	69.64	79.55		
42	56.30	72.87	77.26		
43	59.00	74.16	79.56		
44	61.60	74.55	82.63	4.50	59.69
45	65.90	74.94	87.94	4.74	60.89
46	70.80	77.52	91.33	5.16	63.71
47	78.00	84.75	92.03	6.14	71.97
48	93.10	97.03	95.95	7.04	80.95
49	100.00	100.00	100.00	7.85	91.07
50	102.10	102.84	99.28	8.21	96.55
51	116.20	113.70	102.20	9.26	108.70
52	134.60	116.54	115.50	9.86	110.48
53	141.70	115.50	122.68	10.77	114.76
54	154.20	116.15	132.76	12.85	138.48
55	160.50	116.41	137.88	14.05	152.67
56	169.50	118.09	143.54	14.91	160.18
57	184.00	121.83	151.02	16.11	174.19
58	197.10	125.06	157.60	17.84	194.98
59	204.70	126.49	161.84	18.88	211.82
60	213.50	128.04	166.75	21.32	236.53
verage nnual % ncrease	5.83	3.08	2.66	7.62	5.84

The hospital rate index increased at an annual rate of 5.8%, over 2.6% per annum faster han the CPI over this period. But these entire period rates of growth are only part of ne story. Over the post-depression recovery period 1935-42, increases in hospital rates lagged ehind those of the overall CPI, but hospital rates made up the lost ground by 1945. We nd rapid relative price growth in the hospital sector over the period 1945-54, during which

time average hospital rate growth was 9.9% while the CPI moved ahead at 5.0% per annum. During a period of relatively modest inflation (by today's standards), one finds one-year hospital rates of growth as high as 19.4% (1947-48) and 15.8% (1951-52), the latter coming in a year during which the CPI grew only 2.5%. During the subsequent universal insurance development period (1954-60), when inflation was virtually non-existent (1.6% growth per annum), hospital rates continued to climb about 5.6% per annum.

So much for the history. What is of interest here is the per diem and cost per case behaviour during this period. Per diems clearly exceed rate increases over the period 1932-60, having increased about 1.7% per annum faster than reported hospital rates. In contrast, cost per admission seems, at first glance, to have moved remarkably closely with hospital rates. On more detailed examination, however, neither series mirrors reported rates particularly closely.

The average annual increase in per diems over the period 1932-45 was 4.33%, about 12% slower than the 4.93% average growth for hospital rates over the slightly shorter 1935-45 period. Costs per admission, however, grew only 1.80% per annum, over the period 1932-45, as the patient average length of stay[4] (ALS) fell from 17.5 days in 1935, to 12.8 days in 1945. From 1945-54, the annual rates of increase were 11.7%, 9.6% and 9.9%, respectively, for per diems, costs per case, and hospital rates. Finally, from 1954-60, the corresponding increases were 8.8%, 9.3% and 5.6% respectively. Thus the per diem series may be an acceptable proxy over the period 1932-45, but appears to overstate price growth thereafter. In contrast, costs per admission run behind rate growth until 1945, and mirror rates rather tightly in the fast growth period 1945-54. They may overestimate price increases in the later 1950s run-up to universal insurance, but it should be recalled that the hospital rate series prior to 1960 drops each province as it enters the national insurance program.

The cost per admission series' usefulness as a price index is seriously undermined by the variation in lengths of stay over the pre-insurance period. The per diems embody intertemporal composition problems described below and, in any event, we were able to construct only a partial per diem series (data were not available for 1933-44). While details regarding sources and composition of the "hospital rate" index remain somewhat unclear, it is adopted more by default than design as the most reliable available price series for the pre-national hospital insurance period.

In the national insurance period, hospital rate (or other "price") information is generally unavailable. Keeping in mind that our objective was to estimate as accurately as possible the historical behaviour of service volume and productivity, we turned to aggregate expenditure (cost) and input price data. We describe below, then, the development of a hospital input price, or hospital cost, index for the period 1960-80, which is subsequently grafted onto the earlier hospital rate series, using 1960 as a bridge.

B. A Hospital "Price" Index Under Universal Hospital Insurance

The largest single component of hospital expenditure is the wage bill. Accordingly our major effort in index construction for the post-1960 period was focused on the development of a hospital sector wage index. This was handled in two stages, one for the period 1962-80, the other to develop a bridge back to 1960. The Health and Welfare Canada occasional publication, *Salaries and Wages in Canadian Hospitals*, reports average monthly salaries for the period 1962-81, for anywhere from 13 (in 1962) to 46 (in 1979) classifications of non-medical staff personnel. This series also provides data on the number of full-time employees within each of these classifications, from 1971 to 1981. A number of indexes were constructed from this data base, in order to test the use of alternative base years and subsets of the personnel classification.

If we let

P_{it} = average monthly salary of the ith category of hospital non-medical staff personnel in year t ($i = 1,...,46$; $t = 1962,\ldots,1979$); and

Q_{it} = number of full-time employees in ith category in year t,

then we can first construct a series of Laspeyres wage indexes for the period 1962-81 using alternative base years (t_0),

$$L_t = (\Sigma\, \tilde{P}_{it} Q_{it_o} / \Sigma\, \tilde{P}_{it_o} Q_{it_o}) \times 100$$

$$\text{where } \tilde{P}_{it} = \begin{cases} 0 & \text{if } P_i\text{, 1962 is not available}^5 \\ P_{it} & \text{if } P_i\text{, 1962 is reported.} \end{cases}$$

Since quantity data are available only from 1971, the index was calculated for three base years, 1971, 1975 and 1979. These series appear as L1, L2 and L3 (columns 1 through 3 respectively) in Table 8.

These three initial indexes make two things apparent. First, for the small set of employee categories reported in 1962, the choice of a base year for quantity weights is virtually inconsequential. There was little apparent change over the 1971-79 period in employment mix within these 13 categories. In fact, the lowest inter-year simple correlation among the nine full-time employee vectors (each with 13 elements) was 0.981. Second, the overall growth in wage levels was dramatic. In a span of 19 years, during which the CPI increased 212%, this wage index increased 496-498% (depending on the base year)!

In an attempt to validate this index, which is, after all, constructed using only 13 of a possible 47 personnel categories, we developed three additional Laspeyres indexes over the shorter 1971-81 period. These indexes used 31 (of 46) categories for which data were reported in 1971, all subsequent years missing data categories being subsets of those from 1971. Thus,

$$L_t = (\Sigma\, \tilde{P}_{it} Q_{it_o} / \Sigma \tilde{P}_{it_o} Q_{it_o}) \cdot 100$$

where t runs from 1971 to 1981,

and

$$\tilde{P}_{it} = \begin{cases} 0 & \text{if } P_i\text{, 1971 is not available} \\ P_{it} & \text{if } P_i\text{, 1971 is reported.} \end{cases}$$

TABLE 8. Alternate Hospital Wage Indexes, 1962-81 (1971 = 100)

Year	L1	L2	L3	L4	L5	L6	P1	P2
62	51.363	51.432	51.571					
63								
64								
65								
66	65.255	65.141	65.205					
67	71.596	71.496	71.582					
68	77.525	77.524	77.664					
69	82.646	82.587	82.728					
70	92.441	92.497	92.536					
71	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
72								
73	119.465	119.371	119.277	119.717	119.769	119.672	119.782	119.782
74	142.890	143.078	143.044	142.993	143.323	143.325	143.244	143.244
75	176.803	177.113	177.110	176.459	176.841	176.921	176.841	176.841
76	190.370	190.921	190.902	189.840	190.394	190.476	190.242	190.242
77	208.178	208.063	207.722	207.513	207.489	207.312	207.020	207.020
78								
79	235.494	235.328	234.823	235.272	235.345	235.012	234.984	235.012
80	269.553	270.198	269.963	268.506	269.257	269.246	269.346	269.331
81	306.902	307.685	307.338	305.517	306.508	306.366	307.129	307.079

Sources: See text.

Again this index was computed using three base years, 1971, 1975 and 1979, which generated L4, L5 and L6, respectively, in Table 8.

The Laspeyres indexes L4, L5 and L6 again make clear the irrelevance of choice of base year, suggesting a relatively invariant hospital personnel mix over the larger number of categories and the shorter time period. Even more important from the perspective of attempting to generate a hospital wage index for the post-1960 period, is the fact that the values of L1 through L6 for the 1971-81 period do not differ appreciably. The 13 categories reported in 1962 do an admirable job of proxying wage movements over the more complete set of personnel classes.

Two Paasche indexes (P1 and P2 in Table 8), were also computed for the period 1971-81, as

$$P1_t = (\Sigma \tilde{P}_{it}\tilde{Q}_{it} / \Sigma \tilde{P}_{it_o}\tilde{Q}_{it}) \times 100$$

and

$$P2_t = (\Sigma \tilde{P}_{it}Q_{it} / \Sigma \tilde{P}_{it_o}Q_{it}) \times 100$$

$$\text{where } \tilde{P}_{it} = \begin{cases} 0 & \text{if } P_{it} \text{ is not available for any t,} \\ & t = 1971, 1973\text{-}7, 1979\text{-}81; \\ P_{it} & \text{otherwise.} \end{cases}$$

$$\tilde{Q}_{it} = \begin{cases} 0 & \text{if } Q_{it} \text{ is not available for any t,} \\ & t = 1971, 1973\text{-}7, 1979\text{-}81; \\ Q_{it} & \text{otherwise.} \end{cases}$$

The difference between P1 and P2 is that P1 is computed for each year on a common set of personnel categories, whereas P2 need not be.

As the Paasche Index columns in Table 8 show, not only is there essentially no difference in P1 and P2 but, as one would expect from the relative invariance of personnel mix, there is also little difference between the Paasche and Laspeyres series.

Extension of the 1935-60 hospital rate index requires the use of 1960 as a bridging year. Since the wage bill is such a large component of hospital expenditures, an accurate extrapolation of the 1962-81 wage index back to 1960 was of considerable importance. Data comparable to those used above in the development of the wage index may be available in unpublished form. However, for the purposes of this analysis, we chose to estimate average hourly hospital wages, using published data series, to create the three-year (1960-62) bridge.

If we denote total hospitals' wage bill by WB, total paid-hours by PDH, total hospital expenditure by EXP, the share of total expenditure attributable to wage bill by WSES (wage and salary expenditure share), total patient-days by DAYS, and paid-hours per patient-day by PHPD, then

$$AVWAGE = WB/PDH = (EXP \times WSES)/(DAYS \times PHPD)$$

While we were unable to find a published AVWAGE (average hourly hospital wage) series covering the 1960-62 period, data were available for EXP, WSES, DAYS and PHPD.

Using data on expenditures per day (EXP/DAYS) from Lefebvre, (1976, Table 9); on WSES from Statistics Canada publications 83-215 (Vol. VI, *Hospital Expenditures*, for 1960 and 1961, and unpublished revised data from Statistics Canada for 1962 (described below); and on PHPD from Statistics Canada's *Hospital Statistics, Preliminary Annual Report* series (1965 edition), all referencing PGAS hospitals, an AVWAGE series was constructed for the period 1960-62. These values appear as the beginning of column 2 in Table 9.

Under the assumption that the average hourly wage did not diverge significantly from the wage index over the short period 1960-62, we used the 1962 values of AVWAGE and L1 to bridge the L1 series back to 1960. The complete hospital wage index series appears as column 1 in Table 9.[6]

The special effort required to construct a hospital wage index would, of course, be unnecessary if an alternative were available. We have brought together in Table 9 three series which would be prime candidates as hospital wage index proxies -- hospital average hourly wage, and two average weekly wages and salaries (AVWWS) indexes (industrial composite

and service sector). The hospital average hourly wage represents an extension of our AV WAGE series forward from 1963 to 1980. It was computed in a manner identical to tha for the period 1960-62 described above. "Hospital expenditures per day" (PGAS hospitals) was taken from Lefebvre for the period up to 1974, and subsequent years' data came from Statistics Canada's, *Hospital Statistics - Preliminary Annual Report*, 1980-81.[7] The WSES information for 1962-80 was obtained from the Institutional Statistics Section, Health Divi sion, Statistics Canada. While the *Hospital Statistics* series (83-228) provides WSES infor mation for 1962-75, and the *Preliminary Annual Reports* provide it for subsequent years there are a number of problems with the published series. Commencing in 1969 the wag and salary expenditure share includes wages and salaries paid to medical staff, and com mencing in 1979-80 a separate employee benefits category is broken out. With unpublish ed series detailing expenditure shares for medical remuneration and employee benefits fo the entire 1962-80 period, we attempted to smooth the most acute discontinuities. In par ticular, gross salaries and wages plus employee benefits is our numerator for the WSES series. Medical staff and student remuneration is not included, although we are unclea as to whether employee benefits includes benefits to medical staff. At any rate, the impac would be small. Finally, PHPD was taken from the *Preliminary Annual Report* series (196 edition for 1963-64, 1967 edition for 1965-67, 1975 edition for 1968-75, and yearly edi tions thereafter). The average weekly wages and salaries (AVWWS) indexes come directl from the *Canadian Statistical Review* (Historical Summary, 1970, for the period 1960-70 selected monthly publications thereafter).[8]

The hospital wage index described here increased in value by 500% over the 20-year perio 1960-80, an average annual increase of 9.4%. As Table 9 illustrates, the hospital averag hourly wage increased much more rapidly (737%, or 11.2% per annum), while the industri composite AVWWS (319%, or 7.4% per annum) and the service sector AVWWS (285% or 7.0% per annum) indexes increased at a much slower pace. During the first decade o full universal hospital insurance, the respective annual average rates of increase were 7.5% for our hospital wage index, 9.8% for the hospital average hourly wage, 5.6% for the in dustrial composite AVWWS index, and 5.5% for the service sector AVWWS index. From 1971-75, the respective corresponding average annual increases were 15.3%, 14.5%, 10.2% and 9.9%. From 1975-79 growth rates all slowed somewhat, with the most dramati slowdown showing up in our hospital wage index. The respective annualized growth rate for this period were 7.4%, 11.2%, 9.1% and 7.7%.

TABLE 9. Hospital Labour Price Indexes and Proxies

Year	Hospital Wage Index	Hospital Average Hourly Wage	Hospital Average Hourly Wage Index	Average Weekly Wages and Salaries Index	Average Weekly Wages and Salaries Index Service Sector
60	44.99	1.16	36.07	55.04	55.66
61	48.41	1.25	38.83	56.84	58.72
62	51.36	1.33	41.18	58.51	60.18
63		1.37	42.52	60.50	61.32
64		1.47	45.55	62.85	63.21
65		1.59	49.15	66.12	66.72
66	65.26	1.77	54.88	69.99	71.28
67	71.60	1.95	60.30	74.71	76.49
68	77.53	2.17	67.15	79.83	80.14
69	82.65	2.64	81.72	85.46	85.46
70	92.44	2.95	91.32	92.14	91.97
71	100.00	3.23	100.00	100.00	100.00
72		3.54	109.79	108.41	108.89
73	119.46	3.96	122.59	116.58	116.20
74	142.89	4.81	149.02	129.39	127.95
75	176.80	5.56	172.33	147.73	145.78
76	190.37	6.77	209.73	165.67	162.83
77	208.18	7.37	228.29	181.60	173.78
78		7.91	245.12	192.80	182.63
79	235.49	8.49	263.06	209.42	196.08
80	269.55	9.71	300.90	230.59	214.26

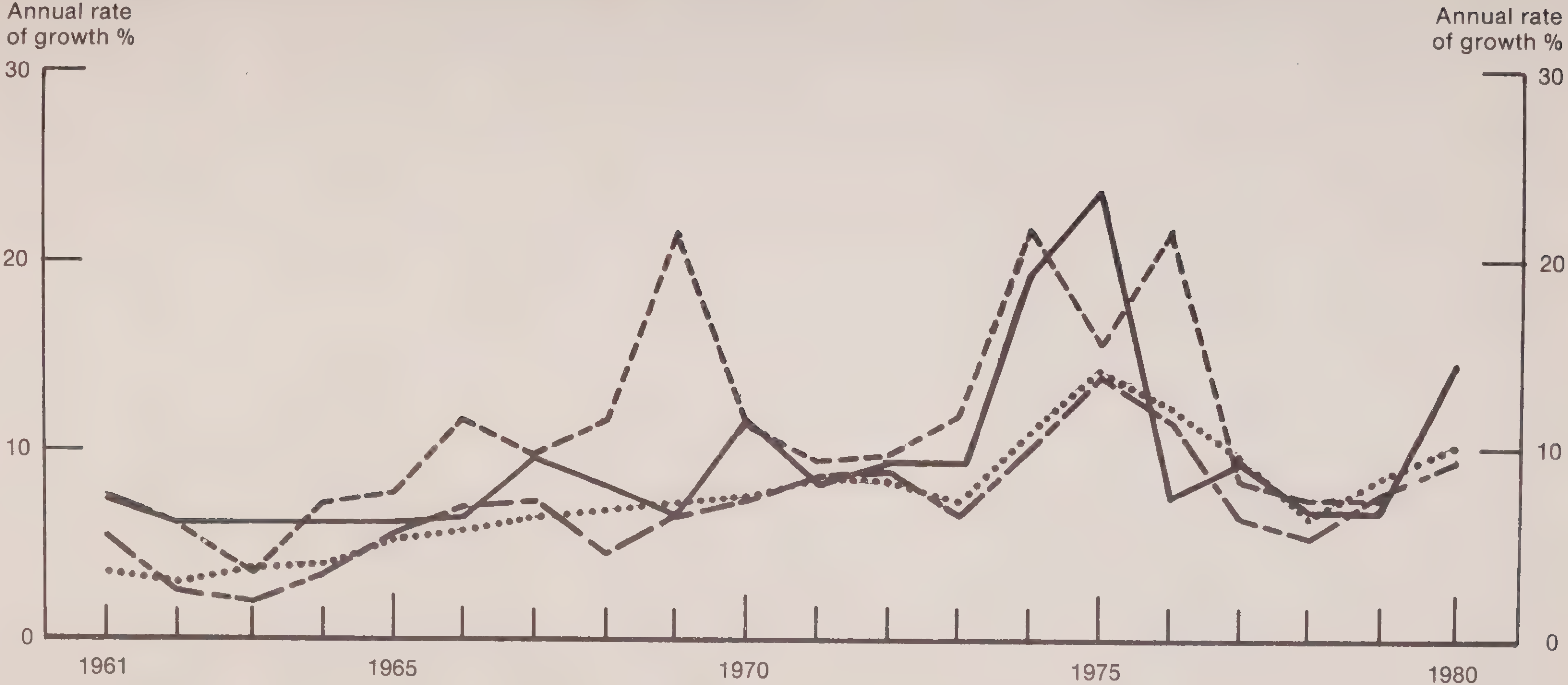
Preliminary and partial data for 1980 and 1981 suggest a resurgence of 1971-75 level growth. Average hospital hourly wages grew 14.4% from 1979 to 1980, compared with 10.1% and 9.3% for AVWWS and its service sub-component. Our hospital wage index has been pushed to 1981 (Table 8), and shows 14.2% per year growth, 1979-81.

Figure 1 shows the annual rates of growth of our hospital wage index and the three comparative indexes. The wage index constructed here (L1) ran ahead of both AVWWS series for most of the 1960-75 period, lagged behind those series during 1975-79, but appears to have forged in front again more recently. The move to public sector restraint in 1982 and beyond may produce another reversal.

Given the relatively invariant employee mix implied by the close comparability of our Paasche and Laspeyres indexes developed earlier, the disparate behaviour of our hospital wage index and the hospital average hourly wage is rather surprising. One might expect factors such as the influence of part-time employment, overtime, different job classification inclusions and the like, to generate some discrepancy. It is, however, difficult to explain differences of the magnitude found here by such factors. Closer scrutiny suggests still unresolved problems with the data. In fact, the movements of our wage index and the average hourly wage over the sub-periods 1960-68 and 1969-80 are not all that different. But our wage index increased 6.6 per cent from 1968 to 1969, while the average hourly wage apparently increased by close to 22%. This discrepancy pointed us back to our construction of the AVWAGE series and, in turn, to the published PHPD series. Paid hours per patient day apparently fell from 14.0 to 13.3 in 1969, enough to account for some but not all of the discrepancy. Whether this was a true fall in labour inputs, or the result of some undocumented shift in hospital or personnel category inclusions has not been determined. In addition, our 'smoothing' of the 1969 discontinuity in the published series' treatment of the medical staff expenditure share may still be incomplete.

Out of this discussion two points seem worth emphasizing. First, while the AVWAGE series is, at first glance, more easily accessible in terms of computation complexity and data availability, intertemporal inconsistencies in the published data which go into its construction make interpretation hazardous and make the series suspect as a reliable labour price index for the Canadian hospital sector. Even with some adjustments using unpublished revised data we seem some way from a reliable AVWAGE series. Second, the even more

Wage Index, Average Weekly Wages and Salaries Composite Index, and Service Sector Average Wages and Salaries Index



Hospital wage index
Hospital average hourly wage index
Average weekly wages and salaries composite index
Service sector average weekly wages and salaries index

tempting (because of availability) proxies represented by the AVWWS indexes grossly underestimate labour price movements in the hospital sector over the entire period 1960-80, and within each of the 1960-71 and 1971-75 sub-periods.

The final stage in the construction of a hospital "price" index for the period 1960-80 consisted of bringing together our wage index with other input price and relative share information. The complete price series is intended here as a "deflator" to allow examination of hospital service volume over time. Since we wish to create a real volume index by converting total operating expenditure in each year (P_tQ_t) to expenditure at base year hospital prices (P_0Q_t), current expenditure must be deflated by a Paasche price index.

If we denote by PI_{it} the price index for input category i in year t, and by ES_{it} the share of total hospital operating expenditure attributable to each input category in year t, then it can be shown (Barer and Evans, 1980) that the appropriate Paasche Index is

$$P_t = 1/[\sum_i (PI_{it_o}/PI_{it})\cdot ES_{it}]$$

The Statistics Canada publications *Hospital Expenditures* (83-215) and *Hospital Statistics* (83-228,83-217) disaggregate total operating expense into four components -- wages and salaries, medical and surgical supplies (MSS), drugs, and other non-departmental supplies and expenses. Because of intertemporal inconsistencies described above, we requested unpublished series for six categories -- medical remuneration, gross salaries and wages, medical and surgical supplies, drugs, employee benefits, and supplies and other expenses.[9] Our GSW (gross salaries and wages) share of expenditure was computed by combining gross salaries and wages with employee benefits, while medical remuneration was added back to supplies and other expenses (although our treatment of employee benefits for that class of personnel is unclear even to us).[10] These adjusted expenditure share series based on unpublished data were extrapolated back from 1962 to 1960 using rates of change in the unpublished expenditure share series over that period.

Unfortunately price series were not available for all four of the resulting expenditure share categories. While we have generated a labour price index suitable for linking with the wages and salaries component, the other input expense categories were more problematic. Statistics Canada does compile and report industry selling price indexes for a

number of product groups, but no such index exists for medical and surgical supplies (MSS). We arbitrarily combined the MSS and drug expense shares, and coupled the resulting share with the Pharmaceuticals and Medicines ISP index. The "price" series chosen for linkage to the residual expense category was the GNE deflator.[11] Both series were taken from the *Canadian Statistical Review*.

Table 10 reports the input price and expenditure share series used to compute the hospital input price index (columns 1 through 6), and the index itself (column 7). The index shows an average annual rate of growth of 7.8% over the 20-year period. This is considerably lower than the 9.4% of the hospital wage index because of lower (assumed) rates of price growth for other inputs. In particular, the price index applied to medical and surgical supplies and drugs, grew at an annual rate of only 2.9% over this period, while the GNE deflator grew an average 5.8% per annum.

The final stage in generating the complete hospital "price" index was simply concatenating the pre- and post-1960 price series, using 1960 as a bridging year. Table 11 reports the complete series in real and nominal terms. The hospital price index increased an average 6.9% per annum over the 45-year period, a period during which the CPI rose at a 4.1% annual clip. The annual growth rates for the pre-1960 period largely reflect our earlier discussions of sub-period price index behaviour in the pre-universal insurance era. From 1960-71, the period covering the development and birth of the provincial universal **medical** insurance plans, hospital prices increased at an annual rate of 6.2%, or 3.3% in real terms. In the immediate post-medical insurance period, 1971-75, hospital prices heated up dramatically, increasing in **real** terms 5.3% per annum, with a one-year 10% real increase from 1974-75. Since 1975, prices have been flat. If we assume that real growth in years with missing data was qualitatively the same as for the respective missing data periods *in toto*, then in only nine of 45 years do we see negative real growth in hospital prices! The period 1950-77 was one of continuous and sustained real growth in the price of hospital services.

With this complete 46-year hospital price index in hand, we are equipped to turn now to examining service volume trends over the same period.

TABLE 10. Hospital Input Price Index, 1960-80

Year	Hospital Wage Index	GSW % of Exp.	Pharm. and Med. ISP Index	MSS and Drugs % of Exp.	GNE Deflator	Supplies and O Exp. % of Exp.	Hospital Input Price Index
60	44.991	0.628	94.527	0.077	72.091	0.294	56.728
61	48.415	0.640	92.764	0.074	72.453	0.288	58.716
62	51.363	0.648	92.208	0.070	73.439	0.281	60.411
63		0.633	92.393	0.070	74.808	0.297	
64		0.640	91.929	0.070	76.663	0.291	
65		0.646	91.837	0.069	79.141	0.285	
66	65.255	0.658	92.857	0.067	82.640	0.275	71.876
67	71.596	0.670	95.455	0.066	85.893	0.264	76.907
68	77.525	0.674	96.568	0.063	88.697	0.262	81.623
69	82.646	0.692	97.959	0.062	92.653	0.245	86.014
70	92.441	0.697	98.794	0.061	96.827	0.242	93.920
71	100.000	0.697	100.000	0.060	100.000	0.243	100.020
72		0.698	102.134	0.058	104.854	0.243	
73	119.465	0.703	102.783	0.057	113.715	0.240	117.124
74	142.890	0.717	109.184	0.054	131.119	0.229	138.396
75	176.803	0.725	120.223	0.054	146.472	0.222	167.143
76	190.370	0.743	126.160	0.050	160.383	0.206	180.829
77	208.178	0.737	130.519	0.052	171.693	0.211	196.480
78		0.731	139.518	0.054	182.307	0.215	
79	235.494	0.729	152.134	0.058	201.333	0.213	223.413
80	269.553	0.729	168.460	0.061	223.714	0.210	253.760

TABLE 11. A Hospital Price Index, 1935-80

Year	Hospital Price Index	CPI	Real Hospital Price Index	Annual Rate of Growth in Real Hospital Price Index
35	12.561	34.783	36.114	
36	12.707	35.457	35.837	−0.766
37	12.901	36.582	35.266	−1.594
38	13.022	36.957	35.237	−0.084
39	13.240	36.657	36.120	2.508
40	13.240	38.156	34.701	−3.929
41	13.434	40.405	33.250	−4.182
42	13.653	42.279	32.292	−2.880
43	14.307	43.028	33.251	2.970
44	14.938	43.253	34.536	3.864
45	15.981	43.478	36.756	6.427
46	17.169	44.978	38.172	3.854
47	18.915	49.175	38.464	0.765
48	22.577	56.297	40.103	4.260
49	24.250	58.021	41.795	4.220
50	24.759	59.670	41.493	−0.722
51	28.178	65.967	42.716	2.946
52	32.640	67.616	48.273	13.010
53	34.362	67.016	51.274	6.217
54	37.393	67.391	55.487	8.216
55	38.921	67.541	57.626	3.855
56	41.104	68.516	59.992	4.105
57	44.620	70.690	63.121	5.216
58	47.797	72.564	65.869	4.353
59	49.640	73.388	67.640	2.689
60	51.774	74.288	69.693	3.036
61	54.199	74.963	72.302	3.742
62	56.537	75.862	74.526	3.077
63		77.211		2.496
64		78.561		2.496
65		80.510		2.496
66	68.687	83.508	82.252	2.496
67	74.347	86.507	85.944	4.488
68	79.484	90.030	88.286	2.726
69	84.895	94.078	90.239	2.212

TABLE 11. A Hospital Price Index, 1935-80 – Concluded

Year	Hospital Price Index	CPI	Real Hospital Price Index	Annual Rate of Growth in Real Hospital Price Index
70	93.468	97.226	96.134	6.533
71	100.000	100.000	100.000	4.021
72		104.798		2.312
73	118.017	112.744	104.677	2.312
74	139.714	125.037	111.738	6.745
75	170.312	138.531	122.942	10.027
76	184.633	148.876	124.019	0.876
77	200.320	160.795	124.581	0.454
78		175.187		–2.408
79	226.900	191.229	118.653	–2.408
80	257.721	210.570	122.392	3.151

The Provision of Hospital Services, 1935-80

While hospital price behaviour, the impetus behind the growth in prices in this secto and price patterns in this relative to other sectors of the economy are all subjects of (a least) academic interest, our major intent in this paper is to use the price series develope in the previous section as a tool with which to examine the other side of the expenditur story -- hospital sector "productivity" and service volume.

We were able to find no published hospital expenditure series for PGAS hospitals coverin the entire period of this analysis. However, data on expenditures per day and on total adul and child patient-days were culled from a variety of sources and used to generate a tota expenditure series. In particular, per diems for the post-1960 period are from Lefebvr (1976) for 1961-74, and from the *Preliminary Annual Report* series thereafter. Total patient days for 1961-65 are from *Hospital Statistics*, 1971: (average daily number of in patients)x365(or 366); for 1966-80 patient-days per 1,000 population (from variou

ᴾ*reliminary Annual Reports*) were combined with population data (from the *Canadian* *Statistical Review*). Columns 1 and 2 in Table 12 report per diems, and patient-days per 1,000 population, for PGAS hospitals. Column 3 represents estimated total hospital ex-)enditure, PGAS hospitals. Our hospital price index is brought forward as column 4. Col-ımns 5, 7, 9 and 11 represent alternate service volume series, while columns 6, 8, 10 and 2 are simply the respective annual rates of change for each volume series. Column 5 (AGG ;VI) is total operating expenses divided by the hospital price index, converted to a base ›f 1971 = 100. This is intended, then, to capture aggregate hospital sector service provi-ion trends. With column 7, in contrast, we examine daily service provision by dividing .verage per diems by the hospital price index. Column 9 represents the per admission/separa-ion service volume series (column 3/(total admissions x column 4)). Total admis-ions/separations are taken from Lefebvre (1976) for 1961-74, with *Preliminary Annual* *?eports* providing the more recent data. Finally, column 11 is service volume per capita :olumn (3)/(population x column 4)).

We noted above that the period 1950-77 saw continuous real growth in hospital prices. :olumn 5 of Table 12 indicates that continuous growth in service volume accompanied ıis relative price growth for most of that period (1952-74). The year in which the largest rice increase occurred (1951-2) was also the only year in the time span 1944-74 in which ıe hospital sector showed a reduction in service provision. The only other year over the ɜriod 1944-80 showing a slowdown in volume was 1974-5, which happens to be the year hich saw the second largest real price increase in the sector (10.0 per cent).

Interestingly the period 1976-79 appears to be unique in this entire 45-year stretch as ıe only period during which stagnant real prices have been accompanied by slow service ɔlume increases. In fact real expenditure over this period was virtually flat (0.8% aggregate ıree-year increase). The only three-year span in which a similar pattern may have existed 1939-42, during which real prices fell over 10%. Unfortunately the absence of yearly ɜr diem data for this period precludes computation of yearly service volume increases, though the average growth over the 1935-44 period was 5.2% per annum.

TABLE 12. Hospital Service Volume Indexes

Year	Per Diem	Days	Total Costs	HPI	AGG SVI	Annual % Change AGG SVI	PDAY SVI	Annual % Change PDAY SVI	PADM SVI	Annual % Change ADM SVI	PCAP SVI	Annual % Change PCAP SVI
35	3.09	1,037.6	34,771.8	12.6	11.0		39.9		60.9		21.9	
36		1,071.0		12.7								
37		1,065.9		12.9								
38		1,076.4		13.0								
39		1,021.6		13.2								
40		1,124.8		13.2		5.2		2.3		−0.7		4.1
41		1,134.7		13.4								
42		1,136.7		13.7								
43		1,184.6		14.3								
44	4.50	1,213.7	65,245.5	14.9	17.3		48.9		57.0		31.3	
45	4.74	1,271.2	72,741.3	16.0	18.1	4.2	48.2	−1.5	55.0	−3.6	32.3	3.1
46	5.16	1,358.2	86,146.3	17.2	19.9	10.2	48.8	1.3	54.4	−1.1	35.0	8.3
47	6.14	1,321.0	101,797.1	18.9	21.4	7.3	52.7	8.0	55.0	1.2	36.7	5.0
48	7.04	1,357.8	122,572.9	22.6	21.6	0.9	50.6	−3.9	52.1	−5.2	36.3	−1.3
49	7.85	1,336.7	141,102.4	24.2	23.1	7.2	52.6	3.8	53.4	2.4	37.0	2.2
50	8.21	1,387.0	156,147.4	24.8	25.0	8.4	53.8	2.4	55.6	4.0	39.4	6.3
51	9.26	1,417.5	183,881.1	28.2	25.9	3.5	53.4	−0.9	54.6	−1.7	39.9	1.3
52	9.86	1,429.0	203,719.5	32.6	24.8	−4.4	49.1	−8.1	48.9	−10.5	37.0	−7.3
53	10.77	1,473.1	235,520.2	34.4	27.2	9.8	50.9	3.8	49.5	1.3	39.5	7.0
54	12.85	1,532.7	301,080.4	37.4	32.0	17.5	55.8	9.6	55.7	12.5	45.1	14.1
55	14.05	1,530.6	337,584.4	38.9	34.4	7.7	58.6	5.0	56.3	0.9	47.3	4.9
56	14.91	1,578.1	378,377.4	41.1	36.5	6.1	58.9	0.5	56.2	−0.1	49.0	3.6
57	16.11	1,578.5	422,386.2	44.6	37.6	2.8	58.6	−0.5	56.0	−0.3	48.8	−0.4
58	17.84	1,624.1	494,875.0	47.8	41.1	9.4	60.6	3.4	58.6	4.6	51.9	6.4
59	18.88	1,649.7	544,531.4	49.6	43.5	5.9	61.8	1.9	60.6	3.5	53.7	3.5
60	21.32	1,643.2	626,040.1	51.8	48.0	10.2	66.9	8.3	65.4	7.8	57.9	7.8

TABLE 12. Hospital Service Volume Indexes – Concluded

Year	Per Diem	Days	Total Costs	HPI	AGG SVI	Annual % Change AGG SVI	PDAY SVI	Annual % Change PDAY SVI	PADM SVI	Annual % Change ADM SVI	PCAP SVI	Annual % Change PCAP SVI
61	23.10	1,639.5	690,717.7	54.2	50.6	5.4	69.2	3.5	67.6	3.5	59.8	3.3
62	24.82	1,721.0	793,776.9	56.5	55.7	10.2	71.3	3.0	72.1	6.6	64.7	8.1
63	26.87	1,753.4	891,912.4									
64	29.18	1,762.4	992,023.6			7.6		4.6		5.0		5.7
65	31.92	1,778.3	1,115,059.0									
66	36.06	1,793.9	1,294,731.0	68.7	74.8		85.3		87.5		80.6	
67	40.38	1,806.2	1,486,256.3	74.3	79.4	6.1	88.2	3.5	91.5	4.6	84.0	4.2
68	45.01	1,850.7	1,724,393.4	79.5	86.1	8.5	92.0	4.3	95.5	4.4	89.7	6.8
69	50.69	1,854.9	1,974,616.5	84.9	92.3	7.2	97.0	5.4	100.0	4.7	94.8	5.7
70	56.24	1,880.2	2,251,996.9	93.5	95.6	3.6	97.7	0.8	99.3	−0.7	96.9	2.1
71	61.58	1,896.6	2,519,100.2	100.0	100.0	4.6	100.0	2.3	100.0	0.7	100.0	3.2
72	68.52	1,858.7	2,782,520.3			3.0		3.0		1.4		1.7
73	77.09	1,849.0	3,153,684.4	118.0	106.1		106.1		102.8		103.4	
74	93.23	1,867.2	3,913,123.1	139.7	111.2	4.8	108.4	2.2	106.2	3.3	106.7	3.2
75	110.30	1,887.7	4,732,063.9	170.3	110.3	−0.8	105.2	−2.9	105.8	−0.4	104.7	−1.9
76	125.79	1,956.9	5,667,798.3	184.6	121.9	10.5	110.6	5.2	118.5	12.0	114.2	9.1
77	135.65	1,985.6	6,270,389.8	200.3	124.3	2.0	110.0	−0.6	121.9	2.9	115.1	0.9
78	147.92	1,980.0	6,880,667.4									
79	162.79	1,902.1	7,338,845.4	226.9	128.4	1.7	116.5	2.9	130.5	3.5	116.8	0.7
80	185.27	1,988.3	8,825,832.9	257.7	135.9	5.9	116.7	0.2	136.0	4.2	122.4	4.7

Per Diem : Total Operating Expense/Total Patient-Days.
Days : Patient-Days per 1,000 Population, Adults and Children, PGAS Hospitals.
Total Costs : Total Operating Expenditure ($000), PGAS Hospitals.
HPI : Hospital Price Index.
AGG SVI : Aggregated Service Volume Index.
PDAY SVI : Per Day Service Volume Index.
PADM SVI : Per Admission Service Volume Index.
PCAP SVI : Per Capita Service Volume Index.

The PDAY index (columns 7 and 8) suggests that daily servicing intensity increased throughout most of the period, but less rapidly than the aggregate service volume index (AGG SVI). The more rapid growth in the aggregate index reflects, of course, the virtually continuous growth in total days of care (only during 1938-39, 1971-72 and 1978-79 did this total not grow). In contrast, the more frequent yearly declines in servicing intensity per admission or separation (PADM SVI) is a result of a dramatic drop in average length of hospital stay. The period 1935-61 saw a fall in ALS from 17.5 days to 11.2 days. Finally, columns 11 and 12 show service volume growth per capita. The PCAP SVI series is of course, simply the aggregate service volume index adjusted for population growth. Since the latter growth has been reasonably stable, this index's high and low growth years tend to reflect the aggregate SVI rather closely.

Table 13 attempts to portray in summary fashion the various hospital trends examined above. The quarter-century 1935-60 was a period of rapid growth in admissions to PGA[S] hospitals and somewhat less sharply declining lengths of stay. The result was an almos[t] 4% per annum growth in aggregate days stay during a period which saw 2% per annum population growth. If we think of service volume as being the product of two component[s] -- hospitalization rates (admissions or days) and servicing intensity (length of stay and in- puts per day for admissions, inputs per day for days) -- then it is clear that over the pre- insurance period, aggregate service volume growth was driven primarily by hospitaliza- tion rates. On an admissions basis, per admission servicing intensity showed essentiall[y] no growth on balance (growth in AGG SVI divided by growth in admissions), althoug[h] servicing intensity did increase with each successive sub-period. Similarly, while the ser- vice volume per day index (PDAY SVI) (servicing intensity) grew at just over two per cen[t] per annum (and considerably faster than that in the latter part of the period), absolu[te] days of care was still the dominant day-related service volume driver.

TABLE 13. Average Annual Rates of Growth (%) in Selected Price and Service Volume Indexes, 1935-80, Canadian PGAS Hospitals

	$ Per Day	Hospital Price Index	Hospital Price Index (constant dollars)	PDAY SVI	Days	ADM/SEP	AGG SVI
1935-45	4.37	2.42	0.16	1.91	3.15	6.18	5.11
1945-54	11.7	9.89	4.67	1.64	4.80	6.37	6.54
1954-60	8.80	5.58	3.88	3.07	3.83	4.21	6.99
1935-60	8.03	5.82	2.65	2.09	3.91	5.77	6.07
1960-71	10.1	6.16	3.33	3.72	3.06	2.84	6.90
1971-75	15.7	14.2	5.30	1.28	1.20	1.05	2.48
1975-80	10.9	8.64	−0.09	2.10	2.12	−0.84	4.26
1960-80	11.4	8.35	2.85	2.82	2.45	1.55	5.34
1935-80	9.52	6.94	2.75	2.41	3.26	3.87	5.75

While the number of days of care was moving along at a brisk clip, per diems were escalating even faster, particularly during the period 1945-54. The 8.03% per annum average increase in per diems over the 25-year period from 1935-60 breaks down as follows: 3.09% for general price inflation, an extra 2.65% for hospital sector-specific real prices, and 2.09% attributable to increased daily servicing intensity. During the 1945-54 period, when per diems were increasing at an annual 11.7% clip, general inflation (5.0%) was the major culprit, followed by additional hospital sector price movement (4.7%) and servicing intensity (1.6%). But in the run-up to universal hospital insurance (1954-60) it was all hospital sector-specific phenomena; inflation cooled off to a modest 1.6% rate, while hospital sector prices tacked an additional 3.9% onto that, and daily servicing intensified at a 3.1% annual rate. In short, the period from 1935 to 1960 was characterized by rapid growth in admissions, a sharp fall in average length of stay, a marked increase in servicing intensity per day during the last six years, and real price increases of a magnitude close to the increases in the CPI.

The decade following the establishment of the hospital insurance plans, during the latter part of which the medical plans also came on-stream, saw continued rapid growth in daily servicing intensity and days of care. For the period 1960-71, per diems increased an average 10.1% per annum, of which the largest component (3.72%) was servicing intensity. Hospital sector real prices accounted for 3.33%, and general inflation for the remaining 2.74%.

The four-year period 1971-75 is intriguing in that it is the period of development of the significant Canada-U.S. gap in per cent of GNP consumed by the health care sector.[14] It is also the period during which the most rapid real hospital price movement and movement in per diems occurred in Canada. The 15.7% annual per diem increase breaks out into 8.49% for inflation, 5.30% for real price movement within the hospital sector, and only 1.28% for servicing intensity. The rapid price escalation was, then, accompanied by a dramatic slowdown in service volume growth. This suggests that the volume side of expenditures was responsible for the divergent Canada-U.S. experience during this period

The most recent five-year period was one in which rates of per diem increase "slowed" back down to a 1960s level (10.9%) but, more significantly, was the first period since the early 1940s during which the hospital sector showed no real price growth. This was a period of rapid general inflation (8.7% per annum), while our hospital price index showed annual growth of 8.6%. Accompanying this stabilization in relative prices was a marked increase in servicing intensity growth relative to that of the 1971-75 period. Even so, servicing intensity during 1975-80 was growing considerably less rapidly than in the 1960-71 era.

The post-universal hospital insurance era 1960-80 shows some marked contrasts to the pre-insurance period: growth in number of admissions slowed sharply and aggregate admissions actually fell during 1975-80; an increase in average length of stay resulted in a 2.5% annual increase in total days of care, but this was also considerably slower growth than in the pre-insurance period; the phenomenal run-up in per diems over this period (11.4% per annum) was about equal parts hospital sector phenomena and general inflation: real hospital prices (2.85%) and servicing intensity (2.82%) together accounted for 5.8%, while the CPI showed annual increases, on average, of 5.3%. Increased servicing intensity, slightly higher hospital sector-specific price growth, and faster inflation resulted in per diems increasing about 2.9 per cent per annum faster during 1960-80 than during the pre-universal insurance quarter-century.

Where does all this leave us? If we treat the analyses in this section as an attempt to clarify historical phenomena within the hospital sector, we might usefully examine the results from the perspective of successive stages of information acquisition. When one begins to think about indicators of hospital sector activity or productivity, one's starting point is usually total expenditures for the sector. That will tell us about growth in nominal, or even real (purchasing power) expenditure, but nothing about relative determinants, productivity or real resource use (all of which, however, happen to be variables of continuing policy interest).

The most obvious and available price/volume split is into days of care and per diems. But perhaps the most significant message coming out of the analyses here is that per diems perform rather differently from "price" movement in this sector, and with good reason. The reason is, of course, integrally linked to the problem of using patient-days as a service volume index, a problem we label the naîve intertemporal homogeneity (NIH) assumption. The use of days to measure service volume amounts to assuming that the input service mix comprising a day of care is intertemporally invariant. Only if growth in our PDAY SV index were (at least close to) zero in all years, could we be comfortable with days as a volume measure and the per diem series as a price indicator. In short:

$$\begin{aligned}\%\Delta\text{TOTEXP} &= (\%\Delta\text{per diem})(\%\Delta\text{days}) \\ &= (\%\Delta\text{price index})(\%\Delta\text{daily servicing intensity})(\%\Delta\text{days})\end{aligned}$$

In the presence of variation in servicing intensity over time, per diems embody both price and quantity information, and days represent only part of the service volume story.

Our service volume indexes contain potentially even more information, however. To this point we have discussed the relative contributions of prices, service volumes (e.g. days or admissions), and servicing intensity, to growth in per diems and total sector costs. Price is a relatively straightforward concept, as is days of care. Servicing intensity is conceptually more problematic.

There are, of course, a number of factors which may contribute to growth over time in servicing intensity, or resource use per day. We deal here with five phenomena -- quality of care, case mix, activity mix, patterns of care, and productivity -- although there may be others. The most straightforward interpretation of the 2.4% average annual increase in resource use per day (an almost 200% aggregate increase over the 45-year period) is what we might term the sophisticated quality fallacy (SQF). Much of the hospital sector research of Martin Feldstein, for example, (see in particular, 1974, 1977) embodies at least an implicit but often an explicit positive relationship (assumed) between servicing intensity (inputs of personnel and supplies per day of care) and quality of care. Within an SQF framework, the PDAY SV index is nothing more than a quality of care index.

If, on the other hand, one takes the more common, and we believe plausible, view that quality must be measured in terms of health outcomes, or improvements in health status attributable to hospital care, the interpretation is much less clear. Available (though unsatisfactory) data on outcomes support the position that more inputs yield better outcomes only very weakly, if at all.[13]

Implicit in an examination of price and volume using per diems and days is the assumption that these measures refer to the same activity base. In fact this is not the case. Per diems are constructed as **total** operating expense divided by days of in-patient care provided. The exercise of dividing per diems by prices to get a daily servicing intensity index carries with it the assumption that the mix of in-patient and other activities has been invariant over the entire period.[14] While Barer and Evans (1980) found inpatient costs to account for a relatively constant (85 per cent) share of operating expenses for B.C. hospitals over the period 1966-73, we have no assurance that this experience can be stretched either intertemporally or geographically. Therefore activity mix must remain as an area for fruitful future research and as a possible partial explanation for the servicing intensity series. About all that can be offered (and with considerable confidence) is that activity mix accounts for at most a small part of the PDAY SVI growth. If anything, non-inpatient hospital activities have grown in importance since 1935, yet as recently as 1973, 85 per cent of operating expense in B.C. hospitals remained attributable to inpatient care.

An increasingly complex case mix is likely a more important contributing factor. While it is far beyond the scope of this paper even to think about quantifying shifts in case mix[15] for the country's hospitals as a whole, we can again draw on Barer and Evans (1980) for some qualitative guidance. Over the seven years of that analysis, servicing intensity for acute in-patient care increased at an annual rate of 3.9%, not unlike the 3.72% found here for the 1960-71 period. Of the 31% aggregate increase over the seven years, the authors estimated from 2 to 5.5% to be attributable to an increasingly complex provincial case mix. Even the upper end of this range implies an average annual rate of case complexity creep of only 0.77%. If this were representative of case mix change over the entire 45-year period of the present analysis, we would still be left with 1.67% per annum (or 110% aggregate) unexplained growth in servicing intensity. But of course that analysis examined only public general hospitals. Shifts in aggregate case mix among PGAS hospitals could have been more dramatic.

A fourth explanation for increasing daily servicing intensity may be shifts in patterns of care. Lengths of stay have fallen considerably since 1935. One might imagine a situation in which patient outcomes are unchanged and aggregate per case servicing is unchanged but packed into fewer days. The result would be growth in our PDAY SVI unrelated to quality of care, case mix or activity mix. In fact, however, the decline in length of stay occurred over the approximate period 1935-54, a period of relatively modest growth in PDAY SVI. Thus, while length of stay fell precipitously (from about 17.5 days to just under 11 days), servicing intensity per admission also fell, so that daily servicing intensity growth only partially compensated for (or explained) the shortened stays. Since 1954, lengths of stay have crept back up to close to 13.4 days.[16] Yet the PDAY SVI has concurrently grown much more rapidly than in the earlier period. Shifting patterns of care are difficult to sustain as explanations for PDAY SVI growth.

To this point, we might safely conclude that shares of the almost 200% increase in daily servicing intensity over this period may be allocated to (i) increases in quality of care (as reflected by improved outcomes), (ii) a slight shift away from in-patient care which has the effect of biasing upward the PDAY SV index; and (iii) an increase in hospital case mix complexity. It is probably also safe to conclude that a significant share of that growth is inexplicable using any of these factors. And that leaves us with a disconcerting residual -- declining productivity.

It appears that on relatively conservative assumptions about quality, activity and case mix, there is a segment of servicing growth attributable to reduced hospital sector productivity. The alternative -- evidence of dramatic increases in outcomes attributable to hospital care -- would be more satisfying. But even sophisticated and complex efforts to compute hospital sector volume indexes do not provide by default the magic link between service volume and hospital sector output. More does not necessarily make patients better than less. If it does not, and case mix changes do not explain the ''more'', then it represents declining productivity.

This brings us full circle to our introductory discussions about the need for precise quantity or service volume data. In the final section we summarize the empirical ground we have covered, and try to put everything back into the context of the CPI.

V. Summary, Research Extensions and Implications

This paper had three major objectives -- consideration of the role of health care in a onsumer price context; the assembly, and creation where necessary, of price series for ospital and physician services in Canada; and the application of these price series as 'deflators'' permitting an examination of real service trends in the hospital and medical are sectors over the period 1935-80.

The peculiar nature of health care and, thus, the manner in which it enters the consumer tility function, argues for its treatment as a public good. This forms a rationale for its xclusion from the CPI, but an equal argument for the availability of price indexes for ther purposes. In particular, with the availability of relatively reliable expenditure data, n examination of real service quantity trends and productivity requires access to equally eliable price information.

Accordingly, our major effort in this paper has been the estimation of improved price ıdexes for medical and hospital care, using wherever possible readily-available, published data. This has served, as a by-product, to highlight some glaring gaps in requisite data nd some easy-to-succumb-to traps in interpretation of existing data. For physician serices the main empirical thrust entailed the development of a methodology for adjusting fficial fee schedule-based price indexes to take account of pre-universal insurance growth ı collection ratios. Prior to the universal public insurance plans, average fees collected y physicians were below official ''list'' fees. But as personal incomes grew and private ısurance spread, the proportion of list fees actually collected was rising. Thus the growth f list fees or reported fees in the pre-Medicare period understated the growth in fees acıally collected. But no hard data were available on the extent of the uncollectibles henomenon, or on its trends during the pre-universal-medical-insurance era. We have arculated and estimated a hypothetical relationship between published and ''real'' medical ervices prices over this period. The output from this estimation exercise was yearly estimates f the proportion of prices unpaid by the uninsured. By combining this information with ata on the proportion of the population with universal coverage, we were able to adjust sted prices to account for growth in the collections ratio.

On the hospital sector side, published price series cease with the establishment of a publi insurance program by the last province, Quebec, in late 1960. Our empirical focus here then, was on creating a pseudo-price index for the period 1961-80. Using Health and Welfar Canada published data on wages and salaries, we first computed a series of labour pric indexes. Since there was little to choose between them, a Laspeyres index based on 1 categories of personnel and 1971 base year personnel weights was adopted. This labou price series became one of three input price series -- the other two being proxies for medica surgical and pharmaceutical supplies, and all other supplies and expenses -- which wer combined with published expenditure share data for the three input classes. The result wa a Paasche hospital "price" index. Finally, this index was bridged at 1960 with the Statistic Canada hospital rate index for the earlier period.

The price series generated for the hospital and medical service sectors were applied t their respective aggregate expenditure series, to generate service volume indexes. On th medical care side, the average 3.2% apparent increase in physician productivity implie by the published medical services fees indexes became a more intuitively reasonable 1.5% after adjustment for the uncollected fees bias in the published price series. As for the hospit sector, we found that a significant segment of growth in per diems over this period wa driven by increases in real resource inputs per day of care. But this increased servicin intensity appears less than fully explained by quality improvements and changes in ca and activity mix. We suggest that only unequivocal evidence of real improvements in p tient health outcomes can head off a conclusion of declining productivity in this secto

The empirical analyses in this paper must be viewed as little more than first tentati probings. There is clearly much more that might fruitfully be done to improve specific tion of the model employed in the medical services analyses, in terms of both better da and more sophisticated modelling. Furthermore, much work remains on the role of changi real and contrived patterns of practice in productivity analyses, and on the manpower pla ning implications of productivity trends. One of the authors is currently attempting to refi one province's fee index by estimating a series of type-of-practice fee indexes. These w eventually allow disaggregation of provincial benefit schedule changes into specialty-specif price movements. Of course the availability of such series would permit examination trends in productivity by specialty.

On the hospital side, the input price index for 1961-80 hinges rather critically on some oft input price data, and some questionable aggregations. For example, we have adopted he GNE deflator as a price series for the "other supplies and expenses" category, and he pharmaceuticals and medicines ISP index for medical and surgical supplies and drugs. n addition, we have only three input categories -- labour, medical and surgical supplies and drugs, and all other. However, some alternative grouping of expenditure categories and price proxies had little impact on our hospital price index. In a similar earlier analysis, Barer and Evans (1980) used data provided by the Statistics Canada *Annual Return of Hospitals* (HS-1 and HS-2) tapes, to disaggregate in-patient expenditure into four categories: abour, medical and surgical supplies plus other supplies and expenses, drugs, and food. This allowed a slightly more precise coupling of expenditure share and price indexes: the pharmaceuticals and specialties ISP index was applied to the drug component, and the ood component (ex. alcoholic beverages) of the CPI was applied to food expenses. The GNE deflator was again the proxy of least resistance for the assorted other expenses. Inerestingly enough, the resulting in-patient hospital price index for B.C. hospitals, 1966-73, ncreased 90%, or at an average annual rate of 9.6%. Over the same seven- year period, he Canadian hospital price index developed in the previous section of this paper showed annual growth of 8.0%. The discrepancy is attributable to some combination of: faster growth in hospital sector input prices in B.C.; faster input price growth in in-patient care han in other hospital activities; different input price proxies used in the two analyses. In any event, the HS-1/HS-2 data base provides the potential for considerable refinement of the price index developed in this paper. There may even be a price series for medical and surgical supplies filed in a dusty corner of Statistics Canada under "unpublished".

These and other refinements would undoubtedly improve the reliability of the price series generated in the analyses of the previous two sections, and in so doing would shed further light on service volume trends and interpretations in hospital and medical care in Canada. But it is unlikely that further empirical and analytical window-dressing will appreciably alter the major qualitative implications of our analyses -- that published medical service ee indexes badly overestimate fees actually received prior to Medicare and, by implicaion, growth in physician productivity; that available proxies such as service sector average weekly wages and salaries, or hospital per diems, do a poor job of proxying "price" trends n the Canadian hospital sector; that there is a significant component of hospital per diem

growth most logically explained by falling hospital sector productivity (in terms of healt outcomes generated by increasing resource inputs). These will, we suspect, stand the te of more rigorous empirical scrutiny.

Going beyond potential improvements and extensions of the analysis herein, we belie that the more accurate measurement of health services prices leads into more general issu of price measurement. As stressed in the introduction, very important issues of health ca planning, delivery, and evaluation turn on the measurement of service prices and quan tities. In general, the quality of present expenditure data is far better than that of the pri or quantity components. Better price data for such a quantitatively and socially signif cant sector of the economy, seems an obvious priority.

But the CPI is not the appropriate framework. Health care is one of a class of service or sectors of the economy, in which consumption or utilization, while apparently carrie out individually, is not really under individual control or free choice in the sense assume in consumer theory. Such goods and services are in every country collectively supplie either directly by government, by quasi-public agencies, or by (or under the regulation o private bodies clothed with public authority and responsibilities.

The consumption of such commodities is not the result of private market transaction and their prices are not determined in competitive or even monopolized markets. Suc prices are the outcome of a negotiation process (which may focus, in fact, on income and an expression of public policy. But the price behaviour of such quasi-public good does matter a great deal, in its welfare implications and for policy design, and it is tim that more comprehensive statistics in this area were prepared. The Statistics Canada pro ject to design a set of National Health Accounts would appear to be a good framewor with which to start.

ootnotes

The strand of economic literature which postulates that consumers can control illness status through their choices of preventive services and other consumption patterns is, unfortunately, without any evidentiary basis.

Series B142 is total patient-days in reporting hospitals. Series B126 and B125 are, respectively, number of hospitals reporting and operating. Then patient-days for this period was estimated as (B142xB125/B126).

Series B141 is number of admissions in reporting hospitals. Admissions is then (B141xB125/B126).

ALS is estimated as patient-days/admissions. It is more accurately patient-days/separations, but separations data were not available and differences are, anyway, most often small.

In theory, one requires $P_{it} = O$ for every t if P_{it} is missing for any t. It turns out, however, that for those categories with salaries reported in 1962, salary information was reported as well in all subsequent years. No data of the P or Q variety were reported for 1963-65, 1972 or 1978.

As noted in Footnote 5 no wage or employment data appear in *Salaries and Wages in Canadian Hospitals* for the years 1963-65, 1972 and 1978.

Commencing in 1977-78, the preliminary annual reports were based on a 1 April to 31 March fiscal year. No adjustment has been made for this inconsistency in the results which follow.

Both indexes are based on the 1960 standard industrial classification.

Even these series contained inconsistencies -- intern and resident salaries appear to have been moved from "supplies and other expenses" to "medical remuneration" from 1969 on; medical remuneration was subsumed within gross salaries and wages in 1980-81; and Quebec's drug and medical and surgical supply expenses were included in supplies and other expenses for 1979-80 and 1980-81. Finally, data were obtained only for the period 1962 to 1981-82.

The medical remuneration share of GSW for 1980-81 was estimated using a linear extrapolation of the medical remuneration share of total expense over the period 1971 to 1979-80. Quebec's drug and medical and surgical supply expenses were "extracted" from supplies and other expenses by applying average shares of those Quebec expenses over the period 1971 to 1977-78 to total Quebec expenditures for 1979-80 and 1980-81.

Recognizing that medical and surgical supply prices may in fact have increased more rapidly than those of drugs, and that prices of supplies and other expenses (embodying food and fuel among other items) likely outstripped growth in the GNE deflator, we attempted some crude sensitivity analysis. When the "supplies and other expenses" share was coupled with the food price index (from the *Canadian Statistical Review*), the MSS share with the GNE deflator, and the drug expense share with the Pharmaceuticals and Medicines ISP index, the resulting hospital price index increased 364% over the 20 years 1960-80 (an average annual increase of 7.97%). As is shown in Table 10, the construction described in the text generated a hospital price index which grew slightly less rapidly (347% aggregate, or an average annual increase of 7.78%). The difference suggests that alternate plausible weightings of the expenditure shares will not appreciably affect the subsequent analysis or implications.

[12] The share in the U.S. moved from 7.7% in 1971 to 8.6% in 1975 (Waldo and Gibson, (1982, p. 19)), while the share in Canada was 7.5% at both ends of the time period (Health and Welfare Canada, (1982a)).

[13] Fraser (1981) has pulled together a number of crude health status indexes for Canada covering varying periods. Infant mortality has dropped off dramatically even since 1960, but accounts for only a small segment of hospital care. Life expectancy rates have shown some upward movement, but again the changes are orders of magnitude apart, and at ''best'' improvement in life expectancy is only partially attributable to hospital care.

[14] Other activites in operating budgets include outpatient and emergency services, education and research, and non-in-patient administration.

[15] Case mix in this context is defined by the distribution of the hospitals' case load across ICD diagnostic categories.

[16] The principal explanation of the rise in lengths of stay appears to be the growth in long-term care utilization in acute and chronic/convalescent hospitals. The length of stay in public general hospitals, not including allied special, has not been rising. But complete series for public general hospitals alone were unavailable in published form.

References

Barer, M.L. and R.G. Evans (1980), ''Hospital Costs Over Time: Approaches to Price Deflation and Output Standardization'', unpublished mimeo, University of British Columbia.

Evans, R.G. (1982), ''The Welfare Economics of Public Health Insurance: Theory and Canadian Practice'', Discussion Paper 82-35, Department of Economics, University of British Columbia.

________ and A.D. Wolfson (1980), ''Faith, Hope and Charity: Health Care in the Utility Function'', Discussion Paper 80-46, Department of Economics, University of British Columbia.

Feldstein, M. (1974), ''The Quality of Hospital Services: An Analysis of Geographic Variation and Intertemporal Change'', in M. Perlman, ed., *The Economics of Health and Medical Care*, pp.402-19, Toronto: Wiley.

________ (1977), ''Quality Change and Demand for Hospital Care'', *Econometrica* 45:7, pp.1681-1702.

Fraser, R.D. (1981), *Commentaries on the Hall Report*, pp.26-88, Toronto: Ontario Economic Council.

Waldo, D.R. and R.M. Gibson (1982), ''National Health Expenditures, 1981'', *Health Care Financing Review* 4:1, pp.1-35.

Primary Data Sources

Health and Welfare Canada (various years), *Canada Health Manpower Inventory*, Healt Information Division, Ottawa.

__________ (1979), *National Health Expenditures in Canada, 1960-1975*, Health Economics and Statistics Division, Ottawa.

__________ (1982a), *National Health Expenditures in Canada, 1970-1979*, preliminary draft, Ottawa.

__________ (1977), *Salaries and Wages in Canadian Hospitals, 1962 to 1975*, Health Economics and Statistics Division, Ottawa.

__________ (1980), *Salaries and Wages in Canadian Hospitals, 1970 to 1977*, Policy, Planning and Information Branch, Ottawa.

__________ (1982b), *Salaries and Wages in Canadian Hospitals, 1979 and 1980*, Policy, Planning and Information Branch, Ottawa.

Irazuzta, J.O. (1979), *A Trend Analysis of Hospital Utilization in Canada*, unpublishe M.Sc. thesis, Department of Health Administration, University of Toronto.

Judek, S. (1964), *Medical Manpower in Canada*, Ottawa: Queen's Printer.

Lefebvre, L.A. (1976), ''Public General and Allied Special Hospitals in Canada'', Healt Division, Statistics Canada, Ottawa.

Statistics Canada (various years), *Canadian Statistical Review*, Ottawa.

__________ (1972), *Canadian Statistical Review - Historical Summary 1970*, catalogu 11-505, Ottawa.

_________ (various years), *Hospital Expenditures*, catalogue 83-215, Vol. VI, Ottawa.

_________ (various years), *Hospital Statistics*, catalogue 83-228, Ottawa.

_________ (various years), *Hospital Statistics, Preliminary Annual Report*, catalogue 83-217, Ottawa.

_________ (various years), *Prices and Prices Indexes*, catalogue 62-002, Ottawa.

Urquhart, M.C. and K.A. Buckley (1965), *Historical Statistics of Canada*, Toronto: Macmillan.

COMMENTS

John S. Greenlees
Office of Prices and Living Conditions
Marilyn E. Manser
Office of Research and Evaluation
U.S. Bureau of Labor Statistics

The major objective of the paper by Barer and Evans (hereafter B-E) is to develop reliabl price and quantity index series for the Canadian health care sector, which is almost ex clusively publicly financed and consequently almost unrepresented in the Consumer Pric Index (CPI). Here we wish to take exception to only one of B-E's arguments, namely tha "given the peculiar nature of health care as a commodity and of its financing mechanism even in the absence of public programs" a CPI is necessarily an inappropriate framewor for health care prices. We feel that this conclusion is reached without sufficient considera tion of the roles of the CPI and of the prices which are used in its computation. Our com ments will be presented in the U.S. context but are generally applicable to private, non compulsory medical care delivery systems.

It is important to note at the beginning that the CPI expenditure weights for physicians services and for hospital and other medical care service charges reflect out-of-pocket ex penses only, and the weight for health insurance reflects only payments made by th employee or consumer (see Ginsburg, 1978). Therefore, what is at issue is whether pric charges for medical care services that are actually paid for out-of-pocket should be in cluded in the index.

B-E begin by asserting that in the presence of extensive employer-provided insuranc "health expenditure has already left the private marketplace." That is, even under a privat insurance system health care costs should be viewed as equivalent to direct taxes, and healt care services should be treated in the same way as roads, public schools, or other good to which no user charge is attached. We disagree. Their position would appear to be muc too extreme, at least in the U.S. context. For one thing, Bureau of the Census (1981) dat show that, in 1979, 46.2% of all civilian workers in the U.S. did not have private healt insurance subsidized by their employers. Although some of these persons are covered unde their spouses' employer-subsidized plan or by Medicare or Medicaid, many of them ar faced with a choice of purchasing health insurance themselves or being uninsured.

Further, even when an employer-subsidized plan is available to the employee, there is onsiderable choice. Employees typically have the option of choosing not to participate n an employer-subsidized plan, and are always free to purchase additional insurance inlividually if they so desire. Many firms and union contracts offer employees a choice among lternative benefit packages and premium rates. Finally, the variation in medical and denal coverage across firms provides consumers with another dimension of choice. Although he interactions between money wage rates and fringe benefit levels create special problems or the interpretation of medical care prices, this does not appear to justify or necessitate heir exclusion from the CPI.

On more theoretical grounds, B-E argue that health care does not belong in the CPI ecause it ''is not a 'good' in the normal sense''. Rather, health care goods and services which have a negative direct impact on utility) are purchased only because of their value n the production of improved health status. Furthermore, one cannot infer a positive relaionship between consumption and utility; a high level of spending on health care is more ikely to result from sickness or accident than from a relaxed budget constraint. This, again, s a line of argument which, although generally accepted as a characterization of the health are market, does not support B-E's conclusions. Pollak (1978) has shown how the theory f the cost-of-living index can be extended to a household production framework in which 'goods'' (such as health care services) are transformed into ''commodities'' (such as health tatus), the levels of which are the arguments in the household's direct utility function. n particular, he demonstrates that the traditional Laspeyres index of goods prices forms n upper bound on a cost-of-living index corresponding to a base utility level and household roduction technology. Thus, there is no need to exclude health care from the index simpy because health status is unmeasurable or health services often painful.

An analogy can be made with fuels and utilities, which are also not demanded for their wn sakes but rather as inputs to ''commodities'' such as shelter or transportation. As n B-E's medical care example, we would not consider a household in Ottawa better off han its Vancouver counterpart because it consumes more electricity, and suburban comnuters are not usually envied for their high gasoline consumption. However, in the theory f the cost-of-living index there is no assumption that utility and consumption are positively orrelated across households. The purpose of the CPI is to compare different price regimes, ot households with different locations, tastes, or household production technologies.

Finally, B-E argue that the role of the physician in prescribing care, by "breaking theoretical connection between the utility function and observed consumption behaviour", makes inclusion of medical care in the CPI a questionable practice. Again, it seems to us that, to a large extent, consumers determine their own care levels, whether by choosing among physicians, obtaining second opinions, or refusing to seek care at all. Here the analogy can be made with television or automobile repair services, and it seems more reasonable to treat medical care like these other services than as a rationed or exogenously determined consumption item.

In the construction of a CPI, goods and services have traditionally been divided into two categories: those to which a price or user charge is attached, and those which are provided without user charge. The latter group, consisting largely of goods provided by governments, is excluded from the index along with other factors affecting a consumer's environment, his time allocation decisions, etc. (see Gillingham, 1974). B-E fail to demonstrate that medical care services inherently belong in this category; consumers exercise considerable freedom in determining their consumption and expenditure levels. B-E also present no separability argument to theoretically justify the exclusion of medical care from a cost-of-living sub-index (see Pollak, 1975). In summary, then, the authors provide insufficient grounds for deletion of the medical care component, either from the U.S. CPI or from the Canadian CPI prior to the imposition of compulsory public programs.

References

Gillingham, Robert. "A Conceptual Framework for the Consumer Price Index", *Proceedings*, American Statistical Association Business and Economic Statistics Section, 1974.

Ginsburg, Daniel H. "Medical Care in the Consumer Price Index", *Monthly Labor Review*, 101 (August 1978), pp.35-39.

Pollak, Robert A. "Subindexes in the Cost of Living Index", *International Economic Review*, 16 (February 1975), pp.135-150.

________ "Welfare Evaluation and the Cost of Living Index in the Household Production Model", *American Economic Review*, 68 (June 1978), pp.285-299.

U.S. Bureau of the Census. "Characteristics of Households and Persons Receiving Noncash Benefits: 1979". Current Population Reports, Series P-23, No. 110, March 1981.

REGIONAL PRICE INDEXES: THE CANADIAN PRACTICE AND SOME POTENTIAL EXTENSIONS

M. Denny and M. Fuss
University of Toronto

SUMMARY

This paper evaluates the possibilities of developing and using regional consumer price information for comparing regional living costs and measuring regional welfare levels.

We begin by assessing some of the current methodological practices and the availability of information on Canadian regional consumer prices. The available data are extensive but somewhat limited for making interregional comparisons. The evidence suggests that the aggregate **relative** *cost of living in Canada's major urban areas has not changed substantially. Housing, which is poorly covered in this data, is probably the item with the largest significant price differential. We argue for more intensive research on the problems of including housing. This inclusion is necessary.*

The second part of the paper develops a practical theoretical framework for measuring regional differences in welfare using the CPI data base. Building on previous work on interregional productivity differentials, we assess the possibilities for comparing interregional well-being. Our method is an accounting procedure which permits the use of detailed data while avoiding the problems of econometric estimation. While there are some definite limitations to this approach, it does provide an extension to current uses of the data base that would be useful.

RÉSUMÉ

L'indice des prix à la consommation (IPC) est l'une des statistiques économiques les plus utilisées. Dans le présent texte, nous désirons analyser l'expérience canadienne de construction et d'utilisation d'IPC régionaux. L'IPC canadien est construit à partir d'échantillons de prix prélevés au niveau régional. Des IPC régionaux, construits par région, sont

en usage depuis plusieurs années. Ceci permet de comparer les taux de croissance des IPC régionaux mais non les niveaux relatifs.

Les échantillons de prix sont collectés pour des produits qui sont spécifiés selon la région par marque de commerce et genre de magasin, ce qui introduit dans les spécifications des produits un facteur d'hétérogénéité entre les régions. C'est pour cette raison que Statistique Canada a résisté à la mise au point de comparaisons globales de prix à la consommation entre les régions. On produit deux ensembles de comparaisons de prix régionaux, mais le champ d'application est incomplet. En particulier, on ne touche pas le coût du logement.

Les éléments de preuve dont nous disposons laissent entendre que les indices des prix à la consommation régionaux agrégés se sont modifiés de façon fort semblable dans diverses régions du Canada. Par conséquent, les mouvements relatifs des prix qui se sont produits ont été communs à la plupart des régions ou se sont équilibrés entre les produits. Les diverses régions à l'intérieur du Canada ne semblent pas être relativement plus chères pour les consommateurs en 1980 qu'elles ne l'étaient en 1960 ou en 1970. La seule exception est probablement le logement.

Les indications directes sur les niveaux relatifs des prix n'indiquent pas que les prix soient très différents dans les diverses régions à l'exception des coûts élevés à Terre-Neuve.

Malheureusement, les résultats de la comparaison des coûts de la vie relatifs entre les régions à l'aide des données de l'IPC ne sont pas entièrement conformes aux résultats obtenus à partir des comparaisons de prix entre les régions. Nous avons recommandé que cette incohérence apparente soit étudiée.

Les difficultés les plus sérieuses à la comparaison des coûts de la vie relatifs proviennent du défaut de mesurer les coûts du logement. Ces coûts, y compris la valeur du terrain englobent les loyers nets associés à ces emplacements comparativement aux autres. Nous croyons qu'on devrait faire un effort sérieux pour inclure les coûts du logement dans les données de prix à la consommation par région disponible.

La deuxième partie de notre texte est plus théorique. À partir de travaux récents sur l'analyse des différentiels d'efficacité entre les régions, nous avons exploré l'application

de ces techniques à la mesure du bien-être des consommateurs selon les régions. On élabore plusieurs formulations différentes permettant l'évaluation des différentiels de bien-être entre les régions. Ces méthodes sont des méthodes de comptabilité qui évitent la nécessité de recourir à l'analyse économétrique. Ceci exige des postulats théoriques relativement plus forts mais permet l'analyse des données fortement désagrégées dans les cas où le nombre d'observations ne permet pas l'estimation. En particulier, de telles méthodes pourraient être utilisées avec la base de données IPC existante ou en l'étendant. Bien que le texte ne contienne pas d'application empirique, notre recherche future entreprendra de tels travaux.

Introduction

The Canadian Consumer Price Index (CPI) has been officially calculated since 1913. It is a national price index that attempts to capture movements through time in the retail prices of a particular basket of goods. Throughout this long history, samples of price movements have been collected in many urban areas and then combined to produce the national CPI.

Since the commodity price samples are based on urban areas it is not surprising that city CPIs have been published for many years. This paper will concentrate on analyzing the construction and uses of regional CPIs. There are two major sections to the paper. Section A considers the currently produced regional consumer price information. An evaluation is made of the procedures and availability of data for regional CPIs. This section concentrates on the conceptual and empirical problems associated with current Canadian practice. The current regional CPIs are published as separate indexes for each city. They permit a comparison of the rates of growth of prices across cities or through time. They do not permit a direct comparison of regional price levels at a point in time. Statistics Canada produces two other price series, the City Average Retail Prices and the Intercity Retail Price Comparisons which are more interesting for price level comparisons.

The following questions will be discussed. What rationales underlie the current weakness of the data for regional price level comparisons? Is it possible to assess the relative regional price levels and their changes with the current data? How consistent are the alternative sources of information about relative price levels? What additional data should be collected to improve the current regional consumer price data base? The current CPI is of

value for relatively narrow purposes. However, there are other uses for the CPI and for the data collected on regional consumer prices. Section B explores one particular use of these data. The consumer price index is often used implicitly as a welfare indicator. In Section B, we will explore the possibilities of using these data to compare the consumer welfare of different regions in Canada. These developments will build directly on previous work by the authors and others on the evaluation of regional productivity, (see Denny, Fuss and May [1981], Denny and Fuss [1983a, b] and Caves, Christensen and Diewert [1982a, b]).

A. Regional Consumer Price Indexes in Canada

A.1 Consumer Price Indexes in Canada

Economists commonly think of the consumer price index as an approximation to the true cost-of-living index. The latter may be defined in terms of the expenditure function. The expenditure function, $e(p,u)$ has a value equal to the minimum expense of attaining any level of well-being, measured by the utility index, u, at prices, $p = (p_1,\ldots,p_n)$. The true cost of living, TCL, may be written,

$$TCL = e(p^1,u^0)/e(p^0,u^0)$$

which is the ratio of the cost of reaching a reference utility level u^0 at prices p^1 to the cost of reaching this same utility level at prices p^0.

It is difficult to attain information on either the expenditure function itself or its argument u. It is also expensive to collect on a continuous basis all of the price information. For these reasons, statistical agencies commonly use procedures to measure the official CPI that minimize these problems at the expense of some errors.

Statistics Canada uses a Laspeyres index,

$$CPI_t = \sum_i s_{i0} \frac{p_{it}}{p_{i0}} \qquad - \qquad \text{national}$$

$$CPI_{jt} = \sum_i s_{ij0} \frac{p_{ijt}}{p_{ij0}} \qquad - \qquad \text{region j.}$$

In practice, the fixed shares s_{i0} and s_{ij0} used by Statistics Canada often correspond to a different time period than the base year, 0, for the prices p_{i0} and p_{ij0}. This deviation from the definition given here is unimportant for our purposes.

There are many papers in this volume and we will attempt to minimize the potential overlap by commenting on aspects of the current Statistics Canada practices that pertain to the construction of regional price indexes.

Statistics Canada collects prices for specific regional cities. Prices for a small number of commodities are collected and movements in these sample prices are used to represent movements in the prices of a larger number of commodities. The pricing sample is based on commodities that are volume sellers in that city (or region) and prices are collected for transactions in the selected items in the volume retail outlets for that city. The price sample for any city reflects the distribution of local retail store types and the local brand preference for commodities.

There are attempts made to maintain the characteristics of the commodity and retail outlet constant **through time** in a given city. There are no attempts to do this across cities. For this reason, the CPI price sample will contain errors if used to directly measure inter-city price differentials. As we will discuss below, this problem underlies Statistics Canada's decision not to produce regional CPIs across cities.

In practice, Statistics Canada does not maintain complete data on the price levels in individual cities. The monthly price samples are used to produce price changes relative to the previous month. If one wants the actual prices calculated from the monthly samples, rather than the price changes, they cannot be provided at present. For direct comparison of price levels, we must consider two other consumer price series.

For a limited number of commodities (60), prices are collected in 26 cities twice a year. The results are published, in unweighted detail only, as the City Average Retail Prices, in dollars, and are directly comparable across cities. These are primarily food products, with detailed specifications, to ensure that identical goods are being priced in all cities. There are no aggregate commodity indexes calculated for these city prices. We will have little to say about these specific commodity prices since we are concerned with broader measures of commodity prices. For information on specific food commodities, the average retail prices across cities provides valuable information.

The second set of price indexes, the Canadian Intercity Retail Price Comparisons, are currently published annually for 11 Canadian cities. Price information is taken from both the CPI and City Average Retail Price samples to improve the comparability of the commodity definitions across cities. The commodities covered include about two-thirds of the expenditure covered by the CPI. The notable absences are shelter and clothing. The omissions are serious. Neverthless, these data are the most comprehensive available for regional price comparisons.

Statistics Canada, like most national statistical agencies, has chosen to stress the importance of measuring ''pure'' price movements both through time and across space. At the same time, the CPI price samples reflect local average preferences for particular brands and/or product characteristics and the local preference and cost structure for different types of retail outlets. Since this latter choice introduces heterogeneity across space, it conflicts with the ''pure'' price movement philosophy when considering price comparisons across space.

At the national level, the CPI does not measure movements in the prices of a fixed national basket. Rather the fixed baskets, to the extent they exist at all, are at the city level. The national index is the weighted average of a number of fixed city baskets.

The City Average Retail Prices and the Intercity Retail Price Comparisons are attempts to partially rectify the ''pure'' price movement principle and apply it to spatial price comparisons. We say partially because the position taken by Statistics Canada is, implicitly, that the ''pure'' price movement principle cannot be successfully implemented across space.

The Intercity Retail Price Comparison series omits clothing, meals outside the home and shelter. These omissions, particularly of shelter, make the available spatial price comparisons uninformative for many of the general uses to which the public would like to have interspatial prices.

There are several interrelated questions that comprise the basis for the controversies about spatial price indexes in Canada.

1. Do we want to maintain the "pure" price movement principle?
2. If we do, how do we apply it to obtain complete coverage in the interspatial price indexes?
3. If we do not, how do we recommend that Statistics Canada proceed?

We do not believe that the "pure" price movement principle should be maintained. If it is maintained then the coverage of commodities in the regional price indexes should be complete. We do not find convincing the argument that clothing and shelter, for example, are sufficiently different in Quebec, British Columbia and Newfoundland to preclude any price comparisons. Rather we would prefer complete coverage even if in some commodity groups the items being priced either are not identical in all cities and/or are of significantly different importance in the average budgets of different cities. We will discuss this issue more extensively in the section on housing below.

A.2 Regional CPIs: Some Empirical Evidence

In this section, we will investigate the available empirical evidence on regional price differentials. In particular, we will focus on the evidence on relative price levels and changes in different parts of Canada. Is there adequate evidence on the relative living costs in different parts of Canada?

Statistics Canada produces CPIs for many Canadian cities. For 12 cities, Table 1 presents the annual average CPI values in 1979 (1971 = 100) for all commodities and a number of sub-categories. The intercity dispersion in the rate of growth of the all-items CPI is low during this eight-year period, 1971-1979. Prices rose rapidly but they rose quite uniformly in all cities.

Six cities (Halifax, Quebec, Montreal, Ottawa, Toronto and Calgary) had the lowest average price increases. Five other cities (Saint John, Winnipeg, Regina, Edmonton and Vancouver) had only slightly faster average rates of growth in prices (9.7% vs. 9.5%). St. John's has had the fastest rate of growth, 10.5%, over the whole period. This swifter growth was due to well above average price increases in food, housing and health.

For all cities, food prices were rising most swiftly, with housing and health a distant second and third. Clothing, recreation and tobacco and alcohol were the commodity groups with the lowest rates of growth in almost all cities.

While there are variations between individual cities in the observed patterns of inflation, they are dominated by the overall consistency of the general pattern described above. Within each of the major sub-categories, there tends to be more intercity dispersion than in the all-items index.

TABLE 1. Regional CPIs, 1979 (1971 = 100)

	All-items	Food	Housing	Clothing	Transportation
St. John's	200.8	256.2	202.9	152.8	177.4
Halifax	187.4	229.7	185.73	152.3	178.4
Saint John	191.7	236.5	189.8	149.6	177.9
Quebec	188.0	231.6	185.4	143.4	178.3
Montreal	188.9	234.5	176.9	159.5	180.2
Ottawa	188.7	223.8	183.7	175.7	179.1
Toronto	189.8	228.9	184.7	162.1	177.9
Winnipeg	192.7	229.5	198.8	165.8	174.9
Regina	190.5	224.2	189.0	162.6	179.3
Edmonton	192.5	224.3	203.6	169.4	178.3
Calgary	188.9	221.7	195.0	169.1	173.4
Vancouver	191.1	230.8	187.7	161.9	181.8

ΓABLE 1. Regional CPIs, 1979 (1971 = 100) – Concluded

	Health and personal care	Recreation and services	Tobacco and alohol
3t. John's	191.9	155.0	165.4
Halifax	168.7	152.0	158.1
Saint John	178.9	167.8	158.3
Quebec	179.1	159.0	167.2
Montreal	181.1	161.2	171.6
Ottawa	173.9	153.1	172.6
Toronto	191.5	157.5	166.6
Winnipeg	180.4	157.7	164.1
Regina	179.3	175.4	170.9
Edmonton	168.1	163.1	148.5
Calgary	177.5	161.2	153.1
Vancouver	178.9	153.1	167.3

ource: Statistics Canada: *Consumer Prices and Price Indexes.*

The regional CPIs (Table 1) are normalized separately for each city which prevents the irect calculation of relative prices between cities. However, we can ask a more limited uestion. **Given the structure of intercity relative prices in 1971, can we use the results n Table 1 to analyze the changes in the intercity relative price structure**? The practical nswer is yes, although some *caveats* must be noted.

Statistics Canada's judgmental sampling procedures do not imply that the same brand f instant coffee, or any other good, is being priced in all cities during any month. Furermore, the type of clothing and housing that is being priced in Vancouver and Quebec nay be distinctly different due to climatic conditions. In general, there is no precise conol defining the identical characteristics of the items chosen for pricing in different cities.

There is an attempt to control and maintain the identical characteristics of any item beng priced through time in a particular city. Consequently, the rate of change of prices or the value of the city consumer price indexes) in any year should reflect, fairly accurate-, **changes** in the structure of relative intercity prices (compared to a base period). The

unknown relative intercity prices in the base year, 1971, are a combination of "pure" price differences and price differences based on different characteristics of the items that are priced in different cities.

As an approximation, we believe the regional CPIs shown in Table 1 represent changes in the intercity structure of relative prices since 1971. These results suggest that with the exception of St. John's, the overall relative cost of living (CPI) between Canada's urban areas did not change sharply during 1971-1979. To extend this conclusion, we have selected further evidence from a different time period.

TABLE 2. Regional CPIs, 1952 (1939 = 100)

	All items	Food	Rent	Fuel and lighting	Clothing
Halifax	177.3	232.0	126.0	151.2	223.9
Saint John	184.9	233.3	127.1	145.9	230.9
Montreal	192.7	252.7	149.9	143.3	196.6
Toronto	183.0	226.0	154.2	174.9	209.2
Winnipeg	180.3	238.2	134.1	131.5	207.0
Saskatoon	183.3	238.9	132.4	155.4	218.8
Edmonton	179.0	241.6	124.8	121.8	218.7
Vancouver	190.3	244.9	134.6	175.8	221.9

Source: Prices and Price Indexes, 1949-52, Table 27 (D.B.S.).

In Table 2, the regional CPIs for 1952 (1939 = 100) are shown for a smaller sample of cities and commodities groups. The period 1939-1952 encompasses the rapid growth in prices after World War II. Once again, there is evidence that the structure of relative intercity prices did not change massively. Montreal and Vancouver became relatively more costly cities than they were in 1939. At the extreme, Montreal's prices, relative to Halifax's, were 8.7% higher at the end of the period.

During both time periods, there are sharp movements in the relative prices of some of the major sub-groups between particular pairs of cities. For example, the relative price of food increased by 16% in St. John's compared to Calgary during 1971-1979 and the

relative prices of housing services increased by 24% in Toronto relative to Edmonton during 1939-1952. There is very little evidence of large consistent changes in the overall level of intercity relative prices during either of these periods. In the 40 years between 1939-1979, these two sub-periods capture a large proportion of the price increases. Appendix B contains a brief discussion of the results for the complete 40-year time period.

The regional CPIs cannot be used to unveil the relative price levels directly. If the structure of relative prices has changed very little in 40 years, we must turn to other sources for the relative price levels. The evidence from the Intercity Retail Price Comparisons provides the most useful source for this purpose.

A more direct comparison of the intercity relative price levels is available from the periodic direct price comparisons undertaken by Statistics Canada. In Table 3 the relative prices for five commodity sub-groups are shown at October 1979. These commodity groups cover about 55% of the CPI basket expenditure and the commodity characteristics have been standardized across cities to a greater extent than is done in the regional CPI procedures.

Relative prices for these commodity groups do not vary enormously across the cities. Costs seem to be higher in the Atlantic Region and Vancouver and lowest on the Prairies. St. John's, Newfoundland has substantially higher prices than other Canadian cities. This is due to its isolated island location and a high retail sales tax.

Food for home consumption is cheapest in Ontario and Quebec. Prices increase as one moves either east or west.

Household operations are least expensive on the Prairies, with the Maritimes and Vancouver having higher than average prices.

Transportation prices are lowest on the Prairies. Montreal and St. John's have much higher prices.

The Prairies also have relatively inexpensive health and personal care. Ottawa, Toronto and Vancouver are the most expensive.

The tobacco and alcohol commodity group has prices that are strongly influenced by tax policies. The low prices in Edmonton and the high prices in St. John's reflect provincial tax policies.

The periodic collection of direct relative prices between cities permits us to assess the trends through time in the relative price levels. Table 4 presents the data for five cities at six points of time, from five commodity groups.

TABLE 3. Comparative Regional Price Levels (October, 1979)

	Home food	Household operations	Transportation	Health and personal care	Tobacco and alcohol
St. John's	117	107	104	100	115
Halifax	103	101	97	95	104
Saint John	106	102	101	101	104
Montreal	98	101	105	98	105
Ottawa	99	99	100	102	100
Toronto	100	99	100	107	99
Winnipeg	102	94	94	94	99
Regina	105	95	94	92	104
Edmonton	103	97	93	97	87
Vancouver	107	103	99	104	98

Source: Statistics Canada, *Consumer Prices and Price Indexes*.

Using the CPI data, the evidence is consistent with the hypothesis that the aggregate relative consumer price index across cities showed relatively little change over periods such as 1971-79 or 1939-79. Changes were larger for commodity sub-groups but the general tendency of similar experiences in price changes was evident.

The direct evidence from the intercity price comparisons (Table 4) does not contradict this notion. Rather this table highlights, by removing the common trend, the existence of some small but definite shifts in the relative prices over time and suggests that some moderate short-run fluctuations have been transitory. The data in Table 4 are not independent of

TABLE 4. City Price Differentials (Toronto = 100)

	Food at home	Household operations	Transportation	Health and personal care	Tobacco and alcohol
St. John's					
May 1969	112.5	--	--	--	129.5
May 1971	110.1	108.6	105.8	90.2	129.9
Feb. 1975	111.2	--	107.9	96.2	126.0
Sept. 1977	116.3	109.4	106.0	96.1	118.8
Oct. 1979	117.0	108.0	104.0	93.5	116.2
Sept. 1981	113.1	108.0	101.0	91.6	124.0
Halifax					
May 1969	106.3	105.7	101.0	98.0	114.7
May 1971	104.0	108.0	101.0	100.0	114.4
Feb. 1975	103.1	--	100.0	89.5	111.0
Sept. 1977	103.1	110.4	101.0	91.3	105.0
Oct. 1979	103.0	102.0	97.0	88.7	105.0
Sept. 1981	103.0	108.0	101.0	91.6	101.0
Montreal					
May 1969	99.0	105.7	106.9	91.0	105.3
May 1971	100.0	108.6	107.8	92.9	102.1
Feb. 1975	101.0	--	103.0	91.4	101.0
Sept. 1977	101.0	107.3	107.0	93.2	100.0
Oct. 1979	98.0	102.0	105.0	91.6	106.0
Sept. 1981	101.0	100.0	106.0	90.7	104.0
Edmonton					
May 1969	103.1	94.3	97.1	106.0	93.7
May 1971	104.0	96.2	93.2	101.8	100.0
Feb. 1975	103.1	--	94.1	94.3	93.0
Sept. 1977	104.1	104.2	94.0	99.0	90.1
Oct. 1979	103.0	98.0	93.0	90.7	87.8
Sept. 1981	102.0	98.0	90.0	90.7	90.0
Vancouver					
May 1969	105.2	105.7	99.0	107.0	98.9
May 1971	108.1	106.7	99.0	105.4	104.1
Feb. 1975	109.2	--	95.0	98.1	95.0
Sept. 1977	109.2	113.5	99.0	104.9	95.0
Oct. 1979	107.0	104.0	99.0	97.2	98.0
Sept. 1981	105.1	105.0	104.0	97.2	102.0

Source: Statistics Canada, *Prices and Prices Indexes* and *Consumer Prices and Price Indexes*.

the CPI price data. Many of the price samples are identical. The difference is that the city price differentials data (Table 4) use a more standardized commodity specification across cities.

To appraise the consistency of the two sources of data on regional price differentials, we completed an additional calculation. Starting from the measured intercity relative price differentials for October 1979 (Table 3) we derived an implicit set of intercity relative price differentials for May 1971. These **calculated** price differentials for May 1971 equal the measured October 1979 differentials pushed backwards to May 1971 using the measured growth of the CPI component from May 1971 to October 1979. That is, the **calculated** May 1971, intercity price differentials are consistent with the measured CPI growth in prices and the observed October 1979 price differentials.

This calculation will permit us to assess the consistency of the evidence on changes in relative price levels. Table 5, shows the measured (M) and our calculated (C) intercity price differentials for May 1971. It is clear that the two sets of numbers are not identical and there is an interesting pattern to the difference between the two estimates. For four of the five commodity groups, the **measured** price differential tends to be consistently higher than the **calculated** price differential. This is true in 30 out of 36 possible cases, excluding health and personal care. If food is also excluded it is true for 24 out of 27 cases in household operations, transportation and tobacco and alcohol.

The opposite pattern occurs in eight out of the nine cases for the health and personal care commodity group.

We do not think that this is accidental but neither do we have an adequate explanation. The differences in the estimates are not trivial relative to the price differentials themselves in many cases.

The implication is that the Consumer Price Indexes were generally rising faster between 1971-1979 for food at home, household operations, transportation and tobacco and alcoholic beverages than is implied by changes in the Intercity Price Differentials between 1971 and 1979. Health and personal care CPIs were rising more slowly than the increase

TABLE 5. Measured (M) and Calculated (C) Intercity Price Differentials, May 1971 (Toronto = 100)

	Food: at home		Household operations		Transportation		Health and personal		Tobacco and alcohol	
	(M)	(C)	(M)	(C)	(M)	(C)	(M)	(C)	(M)	(C)
St. John's	110.1	105.5	108.6	102.8	105.8	111.4	90.2	93.3	129.5	116.4
Halifax	104.0	105.7	108.6	95.8	101.0	98.2	100.0	101.7	114.4	110.8
Saint John	106.1	103.3	107.6	96.1	103.9	102.6	92.9	102.3	118.6	111.2
Montreal	100.0	97.7	108.6	106.8	107.8	102.9	92.9	96.9	102.1	102.0
Ottawa	102.0	101.5	102.9	99.6	100.0	100.4	103.6	104.6	100.0	98.3
Winnipeg	101.0	103.7	95.2	84.0	97.1	95.9	89.3	92.2	103.1	102.4
Regina	105.1	105.3	96.2	92.2	96.1	94.0	90.2	92.0	104.1	103.5
Edmonton	104.0	103.8	96.2	93.9	93.2	93.2	101.8	104.0	100.0	103.4
Vancouver	108.1	105.6	105.7	105.3	99.0	98.1	107.0	104.5	98.9	97.4

Source: Statistics Canada, *Prices and Price Indexes* and *Consumer Prices and Price Indexes*, various issues.

implied by the difference in the price differentials measured at the beginning and the en of the period.

Statistics Canada does not claim that the changes in the intercity price differentials ove time reflect only price changes. Presumably some of the inconsistencies noted above, relat to differences in the comparison products and the weights used in 1971 and 1979. Still we believe that the inconsistencies require further investigation.

The seven cases that do not fit the pattern do not suggest any explanation. They involv seven different cities and four of the five commodity groups. The magnitudes of th discrepancies are particularly noticeable (a) for St. John's, Nfld., (b) for household opera tions in all of the Maritimes and Winnipeg and (c) for health and personal care and tobac co and alcohol in all three Maritime cities. There seems to be larger discrepancies in th east coast cities.

We were not able to complete a more detailed investigation of the apparent discrepan cies in regional price changes measured by these two indexes. Such an investigation shoul be undertaken for longer time periods and a more detailed commodity list. Statistics Canad may be able to clarify the difference in these results. Some of the discrepancy arises du to the switch from pricing local volume sellers in the city CPIs to a common national iten in the Intercity Retail Price (ICRP) Comparisons. The effects of this source could b calculated by Statistics Canada relatively easily for many commodities.

Even with the possibility of having two different estimates of regional price changes the general results are similar. Large shifts in the relative regional prices have not bee prevalent. While this result may not be surprising, there are strong popular beliefs tha housing is the commodity with the largest regional price differential. In the next section the regional variation in housing prices will be considered.

A.3 Regional Housing Prices

The treatment of owner-occupied housing in the calculation of the CPI has been a source of continued dispute. There are other papers in this volume that directly address these issues. This section will limit itself to a discussion of the evolution of regional housing prices and

the importance of including housing in regional price comparisons.

Table 6 presents evidence on the regional evolution of housing prices measured in two slightly different manners. The first three columns are indexes of the price of new houses, excluding land. The fourth column is the June 1982 value of the owner-occupied shelter component of the CPI. Using either measure, housing prices rose sharply and at very different rates in these cities. If the relative price of many commodity groups has not changed, housing provides one major exception.

The level of regional housing prices is not regularly published by Statistics Canada. For illustrative purposes, Table 7 shows the relative price of houses in 12 Canadian cities. These data are based on estimates provided by the Royal Trust Company. The relative price of housing is quite different in Canadian cities. In 1974, the relative price were very high in Toronto and very low in Saint John and Quebec City. Sharp changes in the relative prices occurred between 1974 and 1982. Housing prices in Halifax, Calgary, Edmonton and Vancouver rose sharply relative to other cities.

The cross-section and time series evidence in these two tables is sufficient to establish the existence of sharply different levels and movements in regional housing prices. The differentials in rental housing although not as large are still greater than the price differentials in most commodity groups.

TABLE 6. New House Prices and the CPI (1971 = 100)

	1977	1979	1982	CPI(June 1982)
Montreal	177.7	228.3	308.4	307.5
Toronto	171.6	179.2	224.3	268.1
Ottawa	171.2	196.0	249.8	271.1
Winnipeg	163.5	223.5	268.6	273.8
Calgary	162.8	299.5	366.3	331.8
Edmonton	172.8	302.5	334.7	325.7

Source: Statistics Canada, *Consumer Prices and Price Indexes.*

TABLE 7. Relative Housing Prices[1] (Toronto = 100)

	Aug. 1, 1974	June 1, 1979	July 1, 1982
St. John's	66.2	63.8	57.6
Halifax	69.2	71.4	97.5
Saint John	57.1	54.4	47.5
Quebec	54.1	64.4	54.4
Montreal	66.8	57.4	53.6
Ottawa	73.7	75.0	75.3
Toronto	100.0	100.0	100.0
Winnipeg	64.7	75.3	65.0
Regina	64.7	72.6	70.8
Edmonton	64.7	106.5	90.9
Calgary	64.7	110.1	104.4
Vancouver	75.7	96.3	108.7

Source: Royal Trust Company.

[1] Prices are estimates for a bungalow with similar characteristics in all cities.

If there are significant regional price level differentials for some commodities, we believ that housing is one of the major commodities involved. Consequently, every effort shou be made to collect information on housing price and rental differentials. It is likely tha these regional housing price differentials are the most important price differentials.

Statistics Canada has not collected regional housing price differentials. This decisio is based on the heterogeneity of housing implying that pure price movements cannot b measured. There is no doubt that housing is heterogeneous. However, its importance i total expenditure combined with the existence of significant regional price differences im ply that an attempt to measure regional housing price differentials should be made Heterogeneity is not predominantly an interregional characteristic but exists for housin as a commodity within any community. An improved version of the type of informatio provided by the Royal Trust Company survey would be very useful.

If resources permit, one could sample prices for many types of owner-occupied and ren-
al accommodation comparable to what is done for food at home. This will reduce the eterogeneity. What makes housing somewhat special is the fixed nature of the existing tock and the land component.

Various attributes of the location will be capitalized into the price of the property. For xample, topography, taxes, climate and local public services may change the market value f a site. It is not possible or desirable to attempt to eliminate differences in local haracteristics that are embedded in property values. Rather, it would be useful to stan- ardize the housing characteristics for lot size and floor space and a small number of other hysical characteristics of the property. Regional price differentials will then reflect dif- erences in regional construction costs and the net capitalized value of any quasi-rents ssociated with one city compared to another.

It would be feasible to extend the new house price series in order to permit the develop- ent of a series on constructon cost differentials. These could be used to separate regional ousing price differentials into a construction and non-construction component.

Unfortunately housing, where the price differentials are important, has been excluded om most official interregional price comparisons. Statistics Canada should reconsider s current practice and collect housing price differentials. How these price differentials ter into the CPI or any regional price index is part of the larger debate discussed by hers in this volume.

The available evidence suggests that excluding housing, the aggregate relative prices facing nsumers in different regions of Canada have remained quite similar. We have three recom- endations. First, housing prices must be included to provide an adequate portrayal of e relative costs of living in different parts of Canada. Second, Statistics Canada should ivestigate the reasons for the divergence in relative price level changes suggested by the PI data compared to the Intercity Retail Price Comparisons. Consistency between these o sources is not necessary, although helpful. More important is an understanding of e source of the differences. Users may then choose which series to use. As part of this ivestigation, Statistics Canada should analyze the importance of choosing local products d store definitions for the measured levels and changes in regional CPI's.

B. Regional Welfare and the CPI Data Base

B.1 Extending the Uses of the CPI Data Base

The collection and analysis of the data base for the consumer price index is an expensive procedure. This section will establish some new methods for analyzing that data base. These methods are index number or accounting methods of analysis. We do not wish to suggest that these methods are superior to methods which require the statistical estimation of parameters. Rather our methods are simple versions of what can be done to extend the analysis of consumer behaviour without resorting to econometric techniques.

The consumer price index and its changes are widely used as indicators of the current level of prices relative to the past and of the rate of inflation. Users are seldom interested in price levels or changes by themselves. Rather, the price index is used to adjust the level of money income or other nominal values, to eliminate the consequences of price changes.

The true cost-of-living index, defined earlier, measures the relative cost of attaining a given standard of living at any two sets of commodity prices. To evaluate the welfare of consumers, the cost of reaching two different welfare levels at a reference price vector may be used as a money metric (MM) welfare measure. Using the expenditure function,

$$MM = e(u^1,p)/e(u^0,p) \tag{1}$$

If we have two observations on prices, p^0 and p^1, either of these can be used as the arbitrary reference vector. In general, we can tie the TCL indexes with the MM welfare measures,

$$\frac{e(p^1,u^1)}{e(p^0,u^0)} = \frac{e(p^1,u^0)}{e(p^0,u^0)} \cdot \frac{e(u^1,p^1)}{e(u^0,p^1)}$$

$$= TCL(u^0) \cdot MM(p^1) = TCL(u^1) \cdot MM(p^0) \tag{2}$$

where the arguments of TCL and MM are the reference utility and price levels respectively.

An alternative form of (2), which will prove useful in our subsequent analysis is

$$\Delta \log e = \log e(p^1,u^1) - \log e(p^0,u^0) = \log TCL(u^0) + \log MM(p^1)$$

$$= \log TCL(u^1) + \log MM(p^0) \qquad (3)$$

It is not our intent to pursue these standard topics in depth. The reader may use the rticles by Pollak and Diewert in this volume and the extensive references they contain. he survey by Sen [1979] provides a broader and less technical discussion of the problems f measuring welfare. The money metric welfare measure is simply a ratio of real income ı the two situations. It is this measure of real income that is being sought.

Suppose we observe a consumer in two different situations (perhaps regions of a coun- y) with different price vectors p^0 and p^1 and utility levels, u^0 and u^1. The behaviour f the consumer can be represented by the expenditure function, e(p,u). Initially we will ssume that the expenditure function is identical in the two situations but this will be relaxed elow. This is always done in measuring the TCL or the CPI.

The true expenditure function is unknown, but we will approximate the logarithm of e expenditure function by a quadratic function in the logarithms of prices and of the rdinal utility level,

$$\log E = G(\log p, \log u) \qquad (4)$$

Since G is quadratic in its logarithmic arguments, using Diewert's Quadratic Lemma iewert [1976]) we can express logarithmic changes in expenditure as

$$\Delta \log E = \log E^1 - \log E^0$$

$$= 1/2 \sum_i \left[\frac{\partial G^1}{\partial \log p_i} + \frac{\partial G^0}{\partial \log p_i}\right] \cdot (\log p_i^1 - \log p_i^0)$$

$$+ 1/2 \left[\frac{\partial \log E^1}{\partial \log u} + \frac{\partial \log E^0}{\partial \log u}\right] \cdot (\log u^1 - \log u^0) \qquad (5)$$

The derivative,

$$\frac{\partial G}{\partial \log p_i} = \frac{\partial E}{\partial p_i} \cdot \frac{p_i}{E} = S_i$$

where S_i is the expenditure share on commodity i. We can rewrite (5) as

$$\Delta \log E = 1/2\, \Sigma\, (S_i^1 + S_i^0) \cdot (\log p_i^1 - \log p_i^0)$$
$$+ \; 1/2 \left[\frac{\partial \log E^1}{\partial \log u} + \frac{\partial \log E^0}{\partial \log u}\right] \cdot [\log u^1 - \log u^0] \qquad (6)$$

where $S_i^0 = S_i(p^0,u^0)$ and $S_i^1 = S_i(p^1,u^1)$.

If the utility level is the same in the two situations ($u^1 = u^0 = u$) then the second tern in (6) equals zero and $S_i^0 = S_i(p^0,u)$, $S_i^1 = S_i(p^1,u)$.

The difference in expenditure may be written as

$$\Delta \log E(\bar{u}) = 1/2\, \Sigma\, (S_i^1 + S_i^0) \cdot (\log p_i^1 - \log p_i^0) \qquad (7)$$

The right-hand side of equation (7) is the Tornqvist discrete approximation to a Divisi price index. Moreover the LHS is simply the logarithm of a quadratic approximation (TCL* to the TCL index. Thus,

$$\text{TCL*} = \exp\,(\log E(\bar{u})) \qquad (8)$$

where TCL $= e(p^1,\bar{u})/e(p^0,\bar{u})$. (See equation (2).)

If the utility levels are not identical, it is not possible to use (7) to measure the TCL except in a special case. If preferences can be represented by the Translog expenditure function,

$$E = \alpha_0 + \sum_i \alpha_i p_i + \sum_i \sum_j \gamma_{ij} \log p_i \log p_j$$

$$+ \alpha_u \log u + 1/2\gamma_{uu} (\log u)^2 + \sum_i \gamma_{iu} \log p_i \log u$$

then Diewert [1976] has shown that (7) does provide an exact measure of the TCL for the reference vector $u^* = (u^0 \cdot u^1)^{1/2}$. One does not know anything about u* beyond this relation to the actual unobserved utility levels.

Returning to the general case, (6), we can rewrite this as

$$\Delta\log E = 1/2 \sum_i (S_i^1 + S_i^0) \cdot (\log p_i^1 - \log p_i^0) + 1/2 (\epsilon_{Eu}^1 + \epsilon_{Eu}^1)(\log u^1 - \log u^0) \qquad (10)$$

where ε_{Eu} is the elasticity of expenditure with respect to utility and equals the inverse of the elasticity of utility with respect to income.

This equation states that differences in the logarithm of expenditure in the two situations equal the sum of differences in the price levels holding utility constant, (the first term in (10)) and differences in the logarithms of the utility levels, holding prices constant (the second term).

In contrast with Diewert's result, equation (10) states that no matter what the utility levels are in the two regions, provided they are equal, the difference in the TCL may be approximated by the Tornqvist index, the first term in (10). This is obvious from a comparison of (10) and (3). There is no conflict between our result and Diewert's, simply a choice of interpretations.

Given observations on prices and quantities, the last term in (10) may be calculated residually as

$$\theta(p^1,p^0,u^1,u^0) \equiv 1/2(\epsilon^1_{Eu} + \epsilon^0_{Eu}) \cdot (\log u^1 - \log u^0)$$

$$= [\log E^1 - \log E^0] - 1/2 \sum_i (S^1_i + S^0_i)(\log p^1_i - \log p^0_i) \qquad (11)$$

This is a form of ''money metric'' (c.f. equation (3)) since the difference in the utility levels is being converted into dollars via the elasticities of expenditure with respect to utility.

This is a very useful representation for purposes of interpretation. The money metric $\theta(p^1,p^0,u^1,u^0)$ measures the difference in the welfare between the two situations. As one would expect the RHS states that the money metric, θ, is a real income measure that depends on differences in nominal income (expenditure) $[\log E^1 - \log E^0]$ minus the differences in the prices, $[1/2 \sum_i(S^1_i + S^0_i) \cdot (\log p^1_i - \log p^0_i)]$ (a measure of the cost of living).

Note that this measure of the cost of living is not a TCL measure unless the underlying expenditure function is homothetic or translog, or the utility levels are identical (and hence $\theta = 0$). There is no way to avoid this problem without resorting to econometric estimation Since cost-of-living comparisons will continue to be made using non-econometric price in dices, the Tornqvist index derived above is an appropriate approximation for such purposes

The money metric could be used with a CPI data base and our approximate cost-of living measures to establish indicators of regional welfare differentials. Statistics Canada' current data base would provide a beginning although it suffers from two deficiencies There are only periodic measurements of the expenditure shares. Consequently, one coul either use the Star and Hall [1976] interpretation of the results as an approximation o derive equation (11) for the case of fixed weights. Second, the current price collection pro cedures do not maintain complete records on price levels. As mentioned earlier, there ar regional price level data for an extensive sub-set of all commodities.

This section has developed a basic framework for analyzing regional welfare levels withi a consistent theoretical framework. This framework has substantial limitations but thes are not unique to this approach alone. The major advantage of this methodology is th it provides a basis for non-econometric measures that are theoretically sounder than thos currently available.

A Direct Approach

One can apply the framework discussed above, to analyze differences in welfare starting with the utility function. Assume that the logarithm of the utility function can be approximated by a quadratic function in the logarithms of the quantities of the commodities.

$$\log u = F(\log X) \quad ; \qquad X = (X_1, \ldots X_n)$$

The difference in the utility level in any two situations may be written

$$\log (u^1) - \log (u^0) = 1/2 \sum_i \left[\frac{\partial F^1}{\partial \log X_i} + \frac{\partial F^0}{\partial \log X_i} \right] (\log X_i^1 - \log X_i^0)$$

$$= 1/2 \sum_i [\epsilon_{uy}^1 \cdot S_i^1 + \epsilon_{uy}^0 \cdot S_i^0] \cdot [\log X_i^1 - \log X_i^0] \tag{12}$$

where ε_{uy}^1 is the elasticity of utility with respect to income, y.

If the elasticity, ε_{uy}, were constant then we may rewrite (12) as

$$(\log u^1 - \log u^0)/\epsilon_{uy} = 1/2 \sum_i (S_i^1 + S_i^0)(\log X_i^1 - \log X_i^0) \tag{13}$$

The LHS of (13) is a money metric which can be calculated since all the variables on the RHS side can be calculated. This is a close approximation to the money metric defined as a residual in (11). This follows from the fact that

$$\log E^1 - \log E^0 \cong 1/2 \sum_i (S_i^1 + S_i^0)(\log p_i^1 - \log p_i^0)$$

$$+ 1/2 \sum_i (S_i^1 + S_i^0) \cdot (\log X_i^1 - \log X_i^0). \tag{14}$$

There is a clear advantage to using the expenditure approach in practical applications. The direct utility case requires the assumption that the income elasticity of utility is constant in order to construct an indicator that can be measured without econometric estimation.

B.2 Regional Variations in Expenditure Functions

In comparing expenditure levels, the expenditure function has been assumed to be identical across comparisons. One of the motivations of our earlier work on productivity comparisons was to interpret cases in which the production functions were not identical in th two situations (Denny, Fuss and May [1981], Denny and Fuss [1983a,b]). It should be state that the consumer case is more difficult than the producer one. Since we can measure quantity produced but not utility, additional information must be obtained in the consumer case

Consider the expenditure function for a consumer in region i at time t,

$$e_{it} = G_i(p_{it}, u_{it}, T_{it}) \tag{15}$$

where p_{it} is the vector of prices in region i at time t and T_{it} is an index of time. The logarithm of the expenditure function will be approximated by a quadratic function in the logarithm of p_{it}, u_{it}, T_{it} and D, a vector of dummy variables,

$$\log E_{it} = G(\log p_{it}, \log u_{it}, D, T_{it}). \tag{16}$$

Applying the quadratic lemma to (16) yields

$$\Delta \log E = \log E_{is} - \log E_{Ot}$$

$$= 1/2 \left[\frac{\partial G^i}{\partial D_i} + \frac{\partial G^0}{\partial D_0}\right] \cdot [D_i - D_0]$$

$$+ 1/2 \sum_k \left[\frac{\partial G^{is}}{\partial \log p_k} + \frac{\partial G^{0t}}{\partial \log p_k}\right] \cdot [\log p_k^{is} - \log p_k^{0t}]$$

$$+ 1/2[\epsilon_{uy}^{is} + \epsilon_{uy}^{0t}] \cdot (\log u^{is} - \log u^{0t})$$

$$+ 1/2 \left[\frac{\partial G^{is}}{\partial \log T} + \frac{\partial G^{0t}}{\partial \log T}\right] \cdot [\log T_{is} - \log T_{0t}] \tag{17}$$

where $D_0 = 0$.

The difference in the log of expenditure in the regions at different points of time is due to differences in the regional expenditure functions at any moment of time (the first term,[1] differences in the regional prices (the second term), differences in the regional utility levels (the third term), and differences in the regional expenditure functions through time (perhaps due to changing tastes).

Only the LHS and the second term on the RHS can be calculated from commodity prices and quantities alone. (Using econometric estimation one could estimate the unknown elasticities in the second and fourth terms.) In spite of this problem, the decomposition indicated by (17) does provide a useful method of integrating several special cases to which we now turn.

If all the data are drawn from the same region then we may simplify the expression (17) to

$$\Delta \log E = 1/2 \sum_k (S^s{}_k + S^t{}_k) \cdot (\log p^s_k - \log p^t{}_k) + W_{st} + \mu_{st} \tag{18}$$

where W_{st} and μ_{st} are the last two terms in (17). Provided we are willing to assume that W_{st} and μ_{st} are zero then this equation can be used to determine the true cost of living through time for one region or perhaps one city. When W_{st} and μ_{st} are zero, tastes do not change and the unity level is constant over time. Alternatively, as noted earlier the first term on th RHS provides an approximation to cost-of-living changes even when W_{st} and μ_{st} are not zero.

Statistics Canada does calculate intertemporal CPIs for cities but does not use this formula. The shares currently available for CPI calculations are constant. This implies that the utility function is linear instead of quadratic in the logarithms of the variables. However, the assumptions underlying our formula also are embedded in their formula.

If interest focuses on regional price indexes at a moment of time then (17) becomes,

$$\Delta \log E = 1/2 \sum_k (S_k^i + S_k^0) \cdot (\log p_k^i - \log p_k^0) + W_{i0} + R_{i0} \tag{19}$$

where $$W_{i0} = 1/2[\epsilon_{uy}^i + \epsilon_{uy}^0] \cdot (\log u^i - \log u^0)$$

and $$R_{i0} = 1/2 \left[\frac{\partial G^i}{\partial D_i} + \frac{\partial G^0}{\partial D_0}\right][D_i - D_0] \cdot$$

Using data on prices and quantities we can calculate the first term on the RHS and this will be an estimate of regional differences in the cost of living.

Up to this point we have defined a regional cost-of-living index for a single city through time (CL) and between cities (RCL) at a moment of time. The rate of change of these indexes are

$$\dot{CL}_t = 1/2 \sum_k (S_k^t + S_k^{t-1}) \cdot (\log p_k^t - \log p_k^{t-1})$$

$$R\dot{CL}_{ij} = 1/2 \sum_k (S_k^i + S_k^j) \cdot (\log p_k^i - \log p_k^j)$$

These are acceptable cost-of-living indexes that could be calculated from an extended CPI data base.

Any attempt to proceed further and calculate the residual expenditure difference between the regions in real terms after adjusting for prices results in difficulties. The residual equals the sum of R_{i0} and W_{i0}.

If we believe either one of these is zero, then we can identify the residual as either the regional difference in real income resulting from differences in the expenditure functions across regions (R_{i0}) or as the regional money metric (W_{i0}). Neither one of these assumptions is likely to be strictly accurate. One possibility is to estimate R_{i0} from econometric demand analysis. This will allow one to compute W_{i0} as a residual. Alternatively, it would allow one to answer the following question which is a common application of regional CPIs: What would be the required difference in expenditures ($\Delta \log E$), so that a consumer moving from region 0 to region i would be as well off in region i ($\Delta \log u = 0$) at the prevailing

price level differences. The computed differential is a natural candidate to serve as a moving allowance.

C. Conclusion

Our paper has covered a rather wide range of topics and these have been selective. In concluding, we would like to emphasize that the general high quality of the available consumer price data does not preclude improvements. We have tried to identify areas in which current practices can be improved and analyzed methods that will permit more extensive use of the data.

Footnote

[1] These differences could be due to differences in the socio-economic characteristics of the population or differences in weather and other ''situation specific'' variables.

References

Caves, D.W., L. Christensen and W.E. Diewert [1982], "Multilateral Comparison of Output, Input and Productivity Using Superlative Index Numbers", *Economic Journal*, 92, pp.73-86.

__________ [1982b], "The Economic Theory of Index Numbers and the Measurement of Input, Output and Productivity", *Econometrica*, 50, pp.1393-1414.

Denny, M., M. Fuss and D. May [1981], "Intertemporal Changes in Regional Productivity in Canadian Manufacturing", *Canadian Journal of Economics*, 14, pp.390-408.

Denny, M. and M. Fuss[1983a], "A General Approach to Intertemporal and Interspatial Productivity Comparisons", forthcoming in *Journal of Econometrics.*

__________ [1983b], "The Use of Discrete Variables in Superlative Index Number Comparisons", *International Economic Review*, June 1983.

Diewert, W.E. [1976], "Exact and Superlative Index Numbers", *Journal of Econometrics*, 4, pp.115-145.

__________ [1983], "The Theory of the Cost-of-Living Index and the Measurement of Welfare Change", *Price Level Measurement: Proceedings From a Conference Sponsored by Statistics Canada.*

Pollak, R. [1983], " The Theory of the Cost-of-Living Index", *Price Level Measurement: Proceedings From a Conference Sponsored by Statistics Canada.*

Star, S. and R. Hall [1976], "An Approximate Divisia Index of Total Factor Productivity", *Econometrica*, 44, pp.257-264.

Sen, A. [1979], "The Welfare Basis of Real Income Comparisons: A Survey", *Journal of Economic Literature* 17, pp.1-45.

Statistics Canada [1982], *The Consumer Price Index Reference Paper*, Ottawa.

——— *Consumer Prices and Price Indexes*, Ottawa, various issues.

——— *Prices and Price Indexes*, Ottawa, various issues.

APPENDIX A

Regional Weights and the CPI

The CPI for individual cities depends on a weighting matrix derived from the family expenditure surveys. Define the elements of the matrix, EX

$E_{ij} \equiv$ expenditure on commodity i in city j (dollars)

$E_{ij} \equiv E_{ij}/E_j$ – expenditure share of commodity i in city j .

$E_i \equiv \sum_j E_{ij}$ – expenditure on commodity i in all cities

$E \equiv \sum_j E_j = \sum_i E_i$ – total expenditure on all commodities in all cities

$s_i = E_i/E$ – expenditure share of commodity i in total expenditure, all cities.

The regional CPIs are constructed using a Laspeyres index. The CPI for city j, CPI j at time t may be written

$$CPI_{jt} = \sum_i s_{ij} \frac{p_{ijt}}{p_{ij0}}$$

where j indexes city and t indexes time with $t=0$ in 1978.

The weights are fixed but changed periodically and the only information retained is the relative price, (p_{ijt}/p_{ij0}).

APPENDIX B

Long-Run Patterns of Regional Relative Price Changes

In the text we have argued that the structure of relative prices across Canada's cities may not have changed sharply during the post-war period. To provide a small bit of additional evidence, we have compiled an All-items CPI (1939 = 100) for five Canadian cities between 1939-1979. The compilation involves simple linking and renormalization of the series produced by Statistics Canada. Due to this procedure, the weights and commodities vary discretely through time. While it is doubtful that the results would change sharply if a more consistent indexing procedure was chosen, the results should be read as an example.

Over 40 years, the rate of growth of the CPI in the five cities has been remarkably similar, Table B.1. Montreal, Toronto and Vancouver have experienced almost identical average rates of growth of prices over the whole period and within sub-periods. Edmonton and Halifax have had marginally slower rates of growth of consumer prices. Most of the difference occurred prior to 1959 for both cities and prior to 1951 for Halifax.

Given the limitations of the regional CPIs and the linking procedures adopted by us, this suggests that intercity relative consumer prices have not systematically changed over four decades. To the extent that there are regional price differentials, they must be persistent features of our economic geography or they are not captured by these price indexes.

TABLE B.1 Long-Run Changes in All-items Regional CPIs[1], 1939-79

	Halifax	Montreal	Toronto	Edmonton[2]	Vancouver
1939	100.0	100.0	100.0	100.0	100.0
1943	117.0	120.4	116.5	115.3	117.3
1947	132.7	138.1	134.2	132.8	135.8
1951	172.6	191.6	181.2	178.2	186.3
1955	176.8	192.9	186.5	179.9	192.2
1959	193.9	209.4	202.4	193.1	208.5
1963	202.6	219.2	211.4	200.1	215.0
1967	217.6	234.1	235.3	219.1	234.2
1971	250.1	269.0	266.6	251.3	268.0
1975	342.0	371.5	369.8	345.4	375.5
1979	468.5	508.1	506.8	483.2	512.1
Average Inc. (%)	(3.92)	(4.15)	(4.14)	(4.02)	(4.17)

Source: Statistics Canada, *Price and Price Indexes* and *Consumer Prices and Price Indexes*, various issues.

[1] Weights: 1938 expenditure pattern – 1939-51.
1947-48 expenditure pattern – 1951-59.
1957 expenditure pattern – 1963-71.
1974 expenditure pattern – 1975-79.

[2] Calgary-Edmonton prior to 1971.

QUALITY ADJUSTMENT, HEDONICS, AND MODERN EMPIRICAL DEMAND ANALYSIS

Ernst R. Berndt
Faculty of Applied Economics, Sloan School of Management
Massachusetts Institute of Technology

SUMMARY

It is widely believed that quality characteristics embodied in commodities and services affect consumers' satisfactions and thus the structure of consumers' demands. To the extent that consumer price indexes attempt to approximate "true" cost-of-living indexes, construction of CPI measures should incorporate quality changes over time into the price index formulae. The practical issue facing the government statistician therefore concerns how quality characteristics might best be incorporated into index number formulae, while the academic economist is likely to be most worried about how the resulting price index formulae relates to the modern theory of consumer demand. This paper focuses on issues of how and under what conditions quality adjustment can be accomplished in a way that is consistent with modern flexible functional form demand analysis.

I attempt to show how such quality adjustment can be incorporated into flexible functional form demand analysis, discuss how such an approach addresses the rich detail-parameter parsimony trade-off, and outline how the suggested procedures could be implemented into much rich new empirical research. Implications for index number construction are also be noted.

RÉSUMÉ

On estime généralement que la qualité des biens et services influe sur la statisfaction du consommateur et, partant, sur la demande. Dans la mesure où les indices de prix à la consommation tentent d'évaluer les indices "réels" du coût de la vie, l'établissement

de mesures de l'indice des prix à la consommation devrait inclure à l'intérieur des formules d'indices des prix les changements enregistrés avec le temps au niveau de la qualité. Le problème auquel fait face le statisticien à l'emploi du gouvernement, c'est donc de déterminer de quelle façon le mieux tenir compte de la qualité dans le cadre des formules d'indice, par opposition à l'économiste universitaire, plus susceptible, en tant que théoricien, de vouloir déterminer comment relier les formules d'indice des prix qui en résultent aux théories modernes en matière de demandes de biens et services. Le présent document introduit et énonce un cadre théorique à partir duquel la qualité des produits durables et périssables peut être ajustée par l'entremise de l'intégration de l'analyse hédoniste des prix à l'analyse moderne, souple et fonctionnelle de la demande et comment l'ajustement de la qualité peut ainsi être reliée à la théorie économique des nombres d'indices.

J'explique notamment comment tenir compte de l'ajustement de la qualité à l'intérieur d'une analyse de la demande souple et fonctionnelle, comment une telle approche aborde la question de l'équilibre entre l'abondance de renseignements détaillés et la pénurie de paramètres, et énonce comment les procédures suggérées pourraient être mises en oeuvre dans le cadre d'une nouvelle recherche empirique plus détaillée. Les conséquences sur la construction des nombres indices sont également à noter.

1. Introduction

It is widely believed that quality characteristics embodied in commodities and services affect consumers' satisfactions and thus the structure of consumers' demands. To the extent that consumer price indexes attempt to approximate "true" cost-of-living indexes, construction of CPI measures should incorporate quality changes over time into the price index formulae. The practical issue facing the government statistician therefore concerns how quality characteristics might best be incorporated into index number formulae, while the academic economist is likely to be most worried about how the resulting price index formulae relates to the modern theory of consumer demand. This paper focuses on issues of how and under what conditions quality adjustment can be accomplished in a way that is consistent with modern flexible functional form demand analysis.

At the outset, it is worth noting that the issue of quality adjustment in price index construction is an important one. In the late 1930s in the U.S., for example, public policy debates arose over whether General Motors should be required to vary its prices in order to stabilize production volumes and employment levels. As part of its contribution to this debate, in 1938 GM funded a study by A.T. Court of the Automobile Manufacturer's Association to assess the effects of auto price changes on the total volume of auto sales.[1] Court argued that "...Price indexes in gross error have been widely used as the basis for serious, official discussions of policy",[2] and chastised both the auto manufacturers for failing to co-operate with and provide information to the U.S. Bureau of Labor Statistics,[3] and BLS officials for publishing new automobile price indexes that took no explicit account of changes in physical characteristics (apparently unlike the BLS practices at that time for constructing price indexes of trucks and farm tractors).[4]

As a practical alternative method for constructing price indexes for goods with frequently changing characteristics and specifications based on "objective usefulness", Court proposed a technique by which, given historical data on auto models over time, price was regressed on time and the characteristics of models (in his case, horsepower, weight and wheelbase length). The coefficient on the time variable was then interpreted as the change in the price index, holding usefulness constant. Invoking utilitarian notions, Court called his procedure the hedonic technique, and summarized its purpose by stating that "...Hedonic price comparisons are those which recognize the potential contribution of any commodity, a motor car in this instance, to the welfare and happiness of its purchasers and the community."[5] He then noted that "...Prices per vehicle divided by this index of hedonic comfort would yield valid comparisons in the face of changing specifications."[6]

Incidentally, it is of interest to note that while the BLS official new car price index rose 45 percent over the 1925-35 time period, Court's proposed quality-adjusted new car price index dropped approximately 55 percent.[7] Not surprisingly, GM officials used these empirical findings along with other data in arguing that auto manufacturers had already been reducing quality-adjusted prices, and that any further price decreases designed to stabilize employment would likely lead the auto manufacturers to the "brink of insolvency", for the required break-even volume would be much larger than the price-induced increase in demand for new cars.[8]

This brief discussion makes clear that issues of quality adjustment in the construction of price indexes are important, and that for some time now, a body of literature has existed – hedonic price analysis – that attempts to deal with the quality adjustment problem.[9]

Court's suggestions concerning hedonic price analysis received little attention for almost 20 years, and only in the late 1950s did interest in quality adjustment issues re-emerge.[10] Since 1960, however, a very large number of empirical hedonic studies has appeared in the literature, much of it dealing with quality adjustment for durable goods such as autos, houses, trucks, tractors, refrigerators and computers; these and other hedonic studies have been surveyed by Zvi Griliches [1971a].

A potentially related important development of the last 15 years has been the introduction of "flexible" functional forms (e.g., the generalized Leontief, translog and generalized Box-Cox representations)[11] into empirical studies of commodity or input demand analysis. An attractive feature of these functional forms is that their flexibility has significantly facilitated empirical studies of substitution possibilities among commodities without requiring imposition of prior constraints on substitution elasticities. Moreover, W. Erwin Diewert [1976] has linked the specification of such flexible functional forms to the classic index number literature by demonstrating that there is an equivalence between choice of functional form and choice of index number formula. Hence recent developments in demand theory link the construction of price indexes to the estimation of parameters in demand equations.

Although a very large number of empirical studies of commodity and input demands based on flexible functional forms has been published within the last decade, it is noteworthy that this modern empirical demand analysis has virtually ignored the classic issues of quality adjustment. John Muellbauer, for example, has noted, "...It is a curious feature of the empirical literature that apparently no one has integrated the hedonic approach into budget studies. Perhaps this is because the practitioners on the two sides have not realized they are speaking the same language."[12] In this paper, I attempt to provide a bridge between the empirical hedonic literature that addresses quality adjustment, and the empirical flexible functional form demand analysis literature that reports estimates of income and price

elasticities. Clarification of such links will of course have important implications for empirical research and the construction of price indexes.

One possible way of treating various qualities of commodities is simply to classify them as distinct products. In principle, expansion in the number of commodities to a large number is permitted with flexible functional forms, but in practice such expansion is constrained by finite-sized data bases and the fact that the number of parameters to be estimated increases more rapidly than the number of commodities. Hence there appears to be a trade-off involving detailed specification of commodities and parsimony in parameterization.[13]

Researchers have attacked this trade-off problem in a number of ways. Some, like Giora Hanoch [1978], have urged that simpler and more restrictive functional forms be used, while others, such as Edward Hudson and Dale W. Jorgenson [1974], have partitioned the inputs into separable subsets, and then have estimated each subset separately.

An alternative approach involves attempting to deal with heterogeneous commodities by aggregating them into a single "quality adjusted" measure. For example, following a suggestion of Daniel McFadden [1978, p.62], Richard H. Spady and Ann F. Friedlaender [1978] have specified a flexible cost function for the trucking industry having only a single output (ton-miles), but have allowed the "quality" of this output to be affected by environmental, network and behavioral variables such as average size of shipment, average load, average length of haul, and percent of less-than-haul traffic. This reduction of potentially many outputs to a single quality-adjusted measure considerably reduces the number of free parameters to be estimated, and in a number of studies has yielded very satisfying results.[14]

Spady-Friedlaender do not address the issue of how quality adjustment would be done were their output a durable product rather than a flow variable (ton-miles), but their approach is clearly suggestive of novel procedures for quality-adjusting commodities, thereby incorporating additional detail yet not being encumbered with as large an increase in the number of free parameters to be estimated. For example, instead of treating numerous different "diet" and "non-diet" foods as distinct commodities in the estimation of demand elasticities among food and other commodities, one way of reducing the number

of parameters to be estimated would be to develop a procedure by which these commodities could be aggregated into a single commodity called ''food'', but whose quality would be affected by such nutritional characteristics as, for example, serving size, percent fat, percent protein, caloric and sodium content. The demand estimation would then involve joint estimation of aggregate food and quality parameters. Provided that the number of quality attributes were less than the number of distinct food types, the total number of parameters to be estimated would be reduced. Such an approach has obvious advantages.

In this paper I will attempt to show how such quality adjustment can be incorporated into flexible functional form demand analysis, discuss how such an approach addresses the rich detail-parameter parsimony trade-off, and outline how the suggested procedures could be implemented into much rich new empirical research. Implications for index number construction will also be noted.

The plan of this paper is as follows. In Section II, I introduce quality adjustment of non-durable commodities into demand analysis, and relate them to previous literature dealing with the simple and variable repackaging hypotheses. In Section III, I extend the analysis to quality adjustment of durable goods, a non-trivial task since the distinction between stock (asset) and rental prices turns out to be an important one. In Section IV, I discuss and briefly review hedonic price analysis, emphasizing there the importance of market structure to the interpretation of parameter estimates. Then in Section V, I illustrate the potential significance of the synthesis of modern demand with hedonic price analysis by considering empirical implementation using the translog functional form. A number of potential applications are outlined. Finally, in Section VI, I present brief concluding remarks.

II. Utility and Expenditure Functions with Quality-Adjusted Non-durable Commodities

I begin by assuming the existence of a well-behaved continuous and twice-differentiable utility function F,

$$u = F(x;b) \qquad (1)$$

relating the consumer's utility level u during a time period to a positive vector x of n commodity flow quantities utilized during the time period, $x = [x_1, x_2,...,x_n]$, and a positive scalar index of quality b for each of the n commodities, $b = [b_1, b_2,...,b_n]$. Let u be monotonically increasing in x and in b. Each scalar element of the vector b is in turn specified to be a function of relevant physical and/or economic characteristics, e.g., the nutritional content of foods. Denoting these characteristics as z_i, $z_i = [z_{i1}, z_{i2},...,z_{ik}]$, I specify the function

$$b_i = h_i (z_i), \qquad i = 1,...,n. \tag{2}$$

Note that according to (1), the relationship between u and x depends on the quality of the commodities b. Although initially I will focus here on non-durable commodities, in many cases x_i can be interpreted as the quantity service flow yielded by the stock of a durable good over a given time period; the corresponding price would of course be the rental rather than the asset (stock) price.

Suppose further that the consumer can purchase amounts of the n commodities at fixed positive prices p, $p = [p_1, p_2,...,p_n]$, and that the consumer's "income" or expenditure on the n commodities is denoted by y

where $y = \Sigma_i p_i x_i$

Define the consumer's dual expenditure funcion H,

$$c = H (u; p; b) \tag{3}$$

which indicates the minimum cost c of achieving utility level u given that the consumer faces fixed commodity prices p and commodity qualities b. Note that according to (3), the relationship among c, u and p depends on commodity quality b.

W. Erwin Diewert [1978] has pointed out that while discussions such as that above are couched in the language of consumer demand theory, by redefining y as output, and x, p and b as vectors of input quantity, price and quality, respectively, one can re-interpret

the function F in (1) as the primal production function and the function H in (3) as the dual producer cost function.[15] For most of this paper I will refer to (1) as the utility function and (3) as the expenditure function. It should be noted, however, that results are interpretable in either the utility or production function contexts; indeed, Diewert [1978] invokes a neutral terminology and simply calls (3) the cost function.

The introduction of the quality vector b into the primal (1) and dual function (3) is not new (see Muellbauer [1971b,1974a,b], Robert E. Hall [1968] and Lawrence J. Lau [1982]), but merits special attention. I begin with the special case where the vector b is restricted to $b = [1,1,\ldots,1,b_n]$, i.e. where quality changes affect only the n^{th} commodity, a case which has recently been considered by Lau [1982] in the context of production and cost functions. In this instance the primal function (1) reduces to

$$u = F(x_1,x_2,\ldots,x_n,b_n) \tag{4}$$

where b_n is the quality index associated with the quantity x_n. I now provide intuition on the interpretation of the quality index b_n.

Corresponding to each level of b_n, one may solve (4) to obtain the commodity (input) requirement function for x_n:

$$x_n = f(u,x_1,x_2,\ldots,x_{n-1},b_n). \tag{5}$$

According to (5), x_n is the minimum amount of the commodity (input) required to attain utility (output) level u, given x_1, $x_2,\ldots,x_{n-1}$, and b_n. Now compare the different required quantities of x_n corresponding to alternative quality levels b_{no} and b_{n1}:

$$\frac{x_{n0}}{x_{n1}} = \frac{f(u,x_1,x_2,\ldots,x_{n-1},b_{n0})}{f(u,x_1,x_2,\ldots,x_{n-1},b_{n1})}. \tag{6}$$

As Lau has noted, x_{n0}/x_{n1} represents the **conversion ratio** between two different quality levels of the n^{th} commodity. Note that this conversion ratio in (6) generally depends o

u and all of x. As will be seen shortly, specification of quality conversion factors in ratio form (7) has important implications for the multiplicative specification of commodity quality.

Suppose now that one wishes to obtain a quality-adjusted measure of x_n by writing x_{n0} in terms of x_{n1}, that is to say, measure x_{n0} in units of x_{n1} having quality level b_{n1}. The quantity of x_{n0} in terms of its equivalent quantity in x_{n1} units is given by

$$x^*_{n0} = \left[\frac{f(u,x_1,x_2,\ldots,x_{n-1},b_{n0})}{f(u,x_1,x_2,\ldots,x_{n-1},b_{n1})}\right] \cdot x_{n1} = B_{n0} \cdot x_{n1} \qquad (7)$$

where

$$B_{n0} = \left[\frac{f(u,x_1,x_2,\ldots,x_{n-1},b_{n0})}{f(u,x_1,x_2,\ldots,x_{n-1},b_{n1})}\right]. \qquad (8)$$

Note that B_{n0} reflects relative values of b_{n0} and b_{n1}. Next consider the level of u^* that ould be attained with x^*_{n0} units of x_n having quality level B_{n0}, given by

$$u^* = F(x_1,x_2,\ldots,x^*_{n0},B_{n0}) , \qquad (9)$$

nd compare this u^* with

$$u' = F(x_1,x_2,\ldots,x_{n1},b_{n1}) . \qquad (10)$$

au [1982, p.177] has shown that these two levels are precisely equal, i.e. $u^* = u'$. Thus, ot only does one have a way of quality-adjusting a commodity (input) in terms of a stan-ard unit that is consistent with demand theory, but these equivalent units can also be serted into a utility (production) function defined in terms of the standard unit.

Essentially, the task served by the conversion ratios (6) and quality-adjustment (8) is

to standardize the various qualities into a common unit of measurement. A non-trivial feature of this quality adjustment is that up to a factor of proportionality the various quality-rated x_n are constructed to be perfect substitutes for one another. Note, however, that the important proportionality factor B_n can vary with elements in the vector of characteristics z_n (see (7) and (2)) and need not be constant. Specifically, from (6) and (7), it is seen that the various types of (x_n, b_n) pairs are convertible into one another by the multiplication of a (not necessarily constant) scalar-valued function of the characteristics z_n. For example, in the context of a non-durable commodity such as energy, if heat values such as Btu's were used to aggregate or quality adjust different fuels, i.e. if $b_n = h(z_n)$ were a function only of Btu thermal conversion ratios, it would implicitly be assumed that up to a factor of proportionality (the heat rate proportions) there is perfect substitution among the fuels in consumption. Hence, the quality measure B_n can be viewed as an aggregate index of commodity quality based on the components z_n.

Any empirical implementation of this quality-quantity approach requires careful specification of the conversion function (8) – the quality adjustment measures. The simplest case occurs when the conversion function is specified to be independent of u and $x_1, x_2, \ldots, x_{n-1}$, i.e. when the ratio x_{n0}/x_{n1} is independent of u, $x_1, \ldots, x_{n-1}$ for all u, $x_1, x_2, \ldots, x_{n-1}$, b_{n0} and b_n, and depends only on the characteristics z_n. Lau [1982, p.178] has shown that this occurs if and only if the derivatives of the logarithm of the commodity requirement function (5)

$$\frac{\partial \ln f(u, x_1, x_2, \ldots, x_{n-1}, b_n)}{\partial u}, \quad \frac{\partial \ln f(u, x_1, x_2, \ldots, x_{n-1}, b_n)}{\partial x_i} \qquad (11)$$

are independent of b_n, which implies that the commodity requirement function for x_n must have the form

$$x_n = f(u, x_1, x_2, \ldots, x_{n-1}, b_n) = f(u, x_1, x_2, \ldots, x_{n-1}, h_n(z_n)), \qquad (12)$$

or equivalently, the utility (production) function must have the multiplicative commodity (input) augmentation form

$$u = F[x_1,x_2,\ldots,x_{n-1},h_n(z_n)x_n] = F[x_1,x_2,\ldots,x_{n-1},b_n \cdot x_n]. \tag{13}$$

Moreover, assuming expenditure (cost) minimization, in this case the dual expenditure (cost) function has the form (see Lau [1982, pp.180-182]):

$$C = H[u,p_1,p_2,\ldots,p_{n-1},p_n/h_n(z_n)] \tag{14}$$

$$= H[u,p_1,p_2,\ldots,p_{n-1},p_n/b_n].$$

If improvements in, say, the nutritional content of foods increase food quality b_n, then in (13) the quantity of quality-adjusted foods is augmented, while in (14) the quality-adjusted price of food is reduced. Note also that since the quality-adjusted quantity of food $x_n^* = b_n \cdot x_n$ and the quality-adjusted price $p_n^* = p_n/b_n$, it follows that $p_n^* \cdot x_n^* = p_n \cdot x_n$, i.e., price times quantity is invariant to quality measurement.

As has been emphasized above, the conversion ratio has been specified to be independent of u and $x_1,x_2,\ldots,x_{n-1}$. In the theory of production, a classic example of this particular conversion function specification is the representation of constant exponential factor augmenting technical change. For example, Harrod-neutral factor augmenting technical change is typically represented by

$$L_t^* = L_t\, h_L(t) = L_t e^{\lambda_L (t-t_0)} \tag{15}$$

where labor in quality-adjusted or augmented units at time t is written as labor in base-period units, L_t, multiplied by an exponential function of time, where t_0 is the base-period point in time and λ_L is the constant rate of factor augmentation for labor. The corresponding dual representation for prices is

$$P_{Lt}^* = P_{Lt} e^{-\lambda_L (t-t_0)} \tag{16}$$

In such cases the conversion ratio $b_L = h_L(t)$ is a function only of time. Other specifications of this h_n function are also permissible. For example, in the case of labor input,

h_n could be a function of age, sex, educational attainment and experience of the labor force; or, for capital equipment h_n could be a function of the vintage or horsepower capacity, provided of course that h_n always be independent of u and $x_1,\ldots,x_{n-1}$. Note, however, that the traditional Harrod-neutral specification of technical change is simply a special case of the quality adjustment framework presented here.

Following Franklin M. Fisher and Karl Shell [1968], John Muellbauer [1971b, 1972, 1974b, 1975a] has called this case when the conversion or quality aggregation function is independent of u and $x_1,x_2,\ldots,x_{n-1}$ the **simple repackaging hypothesis**; essentially, quality improvement here implies "more of the same". At the risk of confusing the nomenclature and for reasons that will soon become more obvious, I shall call this type of specification of quality conversion ratios **input price-independent quality adjustment**.

Having expressed quality adjustment in terms of multiplicative factor augmentation functions, I now relate the quality conversion specification to the widely-used hedonic price equations. Given the conversion functions in (7), (8) and (14), the prices of the different (x_n,b_n) commodities (inputs) must, under the assumption of cost minimization, be in proportion to their marginal utilities (productivities), i.e. the effective price per unit of the standardized quality x_n must be equalized at the margin, so that

$$\frac{p_{n0}}{b_{n0}} = \frac{p_{n1}}{b_{n1}} = p_n^* \tag{17}$$

where p_n^* is, at a given point in time, a "base price" constant reflecting the price of the standardized unit. Taking logarithms of (17), one obtains the familiar hedonic price equation relating quality-unadjusted prices to a vector of characteristics

$$\ln p_{n1} = \ln p_n^* + \ln b_{n1} \, , \tag{18}$$

which, from (2) becomes

$$\ln p_{n1} = \ln p_n^* + \ln h_{n1}(z_{n1}). \tag{19}$$

Hence the hedonic price equation (19) corresponding to input price-independent quality adjustment converts the characteristics z_{n1} embodied in x_{n1} into "base price" or effective price units, which can then be inserted into the standardized quality cost function (14).

Suppose that in (19), the quality conversion function $\ln h_{n1}(z_{n1})$ took the log-log form

$$\ln h_{n1}(z_n) = \sum_{k=1}^{K} b_{nk} \ln z_{n1,k}$$

so that (19) could be rewritten as

$$\ln p_{n1} = \ln p_n^* + \sum_{k=1}^{K} b_{nk} \ln z_{n1,k} \tag{20}$$

where the b_{nk} are coefficients on the k^{th} characteristic of the n^{th} commodity. In the classic study by Waugh [1929], for example, prices of vegetables sold at Boston's Fanueil Hall area in the 1920s are related to characteristics such as stem length, coloring, stem diameter, etc. Coefficients on these vegetable characteristics are then interpreted as reflecting the shadow values of the characteristics; in this way vegetable prices are quality-adjusted.

In the context of durable goods, if, for example, a cross-section of rental price and characteristic data were available for a number of alternative models of a durable good such as refrigerators, trucks, or farm tractors, regression estimates of the coefficients b_{nk} could be interpreted as estimates of the shadow values or shadow prices of the characteristics used in converting quality variations into a standardized unit. Moreover, following Robert E. Hall [1971, p.264], further interpretation of the entire hedonic price equation (20) can be obtained by expressing each of the $z_{n1,k}$ as ratios of the value of this characteristic in the model under consideration (here, model 1) to its value in, say, the j^{th} model, i.e.,

$$z^*_{n1,k} = \frac{z_{n1,k}}{z_{nj,k}}, \quad k=1,\ldots,K \tag{21}$$

for all models. This corresponds to B_{n0} in (8) being a relative augmentation index. If parameters in equation (20) were then estimated with the $z^*_{n1,k}$ replacing the $z_{n1,k}$ for all models, the intercept term $\ln p^*_n$ could be interpreted as the price index of the standardized j^{th} model; any other model embodying the same characteristics as that in the j^{th} model would have all $z^*_{nk} = 1$, therefore all $\ln z^*_{nk} = 0$, and hence would have the same effective price as the j^{th} model. Models embodying alternative characteristic combinations would of course have different effective or quality-adjusted price indexes relative to the standardized model.

In the previous paragraphs I have considered the case where the conversion ratio x_{n1}/x_{n0} or characteristic aggregation function is independent of u and $x_1,\ldots,x_{n-1}$.[16] It is desirable to relax this condition, since it is highly restrictive; for example, conversion ratios between two air conditioners with differing energy-efficiency ratios (EER's) but of the same size might well depend on the price of electricity, and such a case is not allowed when the conversion function is of the simple repackaging form, i.e. when the quality-adjustment conversion function (8) is price-independent. So let us now relax the previous assumption, and consider the case in which the conversion ratio for the n^{th} commodity is still independent of u but is a function of commodity level x_{n-1}, i.e. $b_n = h_n(x_{n-1},z_n)$. As will be seen, this has important implications.

Specifically, when the conversion function (8) is independent of u and $x_1,x_2,\ldots,x_{n-2}$, the commodity (input) requirement function must have the form (see Lau [1982, p.182]):

$$x_n = f(u,x_1,x_2,\ldots,x_{n-1},b_n) = f(u,x_1,x_2,\ldots,x_{n-1},h_n(z_n,x_{n-1})) \tag{22}$$

which implies that the conversion function is of the form

$$\frac{x_{n0}}{x_{n1}} = \left[\frac{f(u,x_1,x_2,\ldots,x_{n-1},b_{n0})}{f(u,x_1,x_2,\ldots,x_{n-1},b_{n1})}\right] = \left[\frac{h_{n0}(x_{n-1},z_{n0})}{h_{n1}(x_{n-1},z_{n1})}\right], \tag{23}$$

and that the corresponding utility (production) function can be written as

$$u = F(x_1, x_2, \ldots, x_{n-1}, b_n \cdot x_n) \tag{24}$$

$$= F(x_1, x_2, \ldots, x_{n-1}, h_n(x_{n-1}, z_n) \cdot x_n).$$

If two quantities of x_n, say x_{n0} and x_{n1}, are both consumed, then the cost minimization assumption requires that

$$\frac{p_n(b_{n1})}{h_{n1}(x_{n-1}, z_{n1})} = \frac{p_n(b_{n0})}{h_{n0}(x_{n-1}, z_{n0})} = p_n^* \tag{25}$$

which implies the generalized hedonic price equation

$$\ln p_n(b_{n1}) = \ln p_n^* + \ln h_{n1}(x_{n-1}, z_{n1}). \tag{26}$$

Notice that in (26), the hedonic price conversion function h_{n1} depends not only on the characteristics z_{n1}, but also on the quantity x_{n-1}.[17]

In the context of air conditioners, for example, quality-adjusted rental prices could be regressed on characteristics such as Btu output, noise level, and annual operating costs, where annual operating costs depend on the energy-efficiency ratio and the quantity of electricity consumed.

At this point it is worth noting that Lau [1982, p.183] has shown that if one assumes cost minimization and specifies an expenditure (cost) function dual to the utility (production) function, the expenditure (cost) function will be of the form

$$C = H[u, p_1, p_2, \ldots, p_{n-1}, p_n'/h_n(p_{n-1}, z_n)] \tag{27}$$

However, in general unless C is homothetic,

$$h_n'(p_{n-1}, z_n) \neq h_n(x_{n-1}, z_n) \tag{28}$$

Hence, primal and dual conversion factors are not numerically equivalent unless the utility (production) function is homothetic (see Lau [1982, p.183]).

When two or more x_n are utilized at the same set of p_{n-1} as in (27), under cost minimization it must be the case that

$$h_n'(p_{n-1}, z_n) p_n^* = p_n(b_n) \tag{29}$$

so that once more one has an hedonic price equation

$$\ln p_n(b_n) = \ln p_n^* + \ln h_n(p_{n-1}, z_n) \tag{30}$$

which now depends not only on z_n, but also on p_{n-1}. Again, in the context of air conditioners, by (30) quality-unadjusted prices are regressed on a set of characteristics (Btu output, noise level, etc.) and annual operating costs which are of course a function of p_{n-1} – the price of electricity.

Generalizing slightly the analysis of Fisher-Shell [1968], John Muellbauer [1974b, p.8] calls this specification of the conversion function the **variable repackaging hypothesis.**[18] In this more general context, the aggregation of characteristics into a scalar quality index depends on prices of certain commodities or inputs; hence I call it **price-dependent quality adjustment**.

Note that in the context of a durable such as used autos, price-dependent quality adjustment would permit quality adjustment of two autos to depend not only on their physical characteristics (e.g. horsepower, interior space, weight), but also on the price of fuels (such as gasoline and diesel fuel). This suggests that in any empirical analysis one could test the simple versus the variable repackaging hypothesis by testing suitable parameter restrictions using classical hypothesis testing procedures, just as others have done in testing for separability of production or cost functions. Such an exercise would however require careful distinction between rental and asset prices of durable goods. Hence I now turn to a discussion of capital stocks and capital service flows, or alternatively, capital asset (stock) prices and capital rental prices.

II. Quality Adjustment: Extension to Durable Commodities or Inputs

In the context of durable goods, it is of course the case that not only do there exist varia-ions among different models of the same age or vintage with varying characteristics, but here also occur significant efficiency differentials among different ages or vintages of the ame model. While both these differences can be viewed as variations in quality, the latter ave traditionally been termed deterioration differences, a convention I will follow here.

With durable goods, quality adjustment will standardize assets of different vintages and haracteristic combinations into a common unit, i.e. quality adjustment will handle the sue of depreciation. The issue, of course, is what factors affect the quality conversion inction. As will be seen, traditional measures based on constant and equal geometric epreciation rates correspond with a special case of the simple repackaging (price indepen-ent) hypothesis whereas, for example, energy price-induced economic depreciation of iergy-inefficient used autos corresponds with the variable (price dependent) repackaging ypothesis.

Assume that the asset or stock price of the n^{th} capital good of vintage at time t is equal the present value of its future services,

$$q_{n,t,\phi} = \sum_{s=0}^{T_n - \phi} \left(\frac{1}{1 + r}\right)^s V_{n,t+s,\phi+s} \tag{31}$$

iere T_n is the lifetime of the asset, r is the rate of interest (assumed to remain constant er time), and $V_{n,t,\phi}$ is the value (i.e. price times quantity) at time t of the flow of ser-es of the n^{th} capital good of vintage ϕ. Lifetimes and prices are assumed to be fixed d known with certainty.

This value can be decomposed into rental price and quantity flow components in a number different ways. I begin with the simple repackaging (price independent) type of decom-sition, analogous to (17) where $p_{n1} = p_n b_{n1}$ and the b_{n1} are independent of u and $x_{\ldots}, x_{n-1}$. In the present context, consistent with the simple repackaging hypothesis, one 1 specify fixed conversion ratios both between capital services from different ages of

the same model, and between capital services from different models of a given age, sc that deterioration in capital services takes place independently of the year in which th good was produced and of the year in which services are used. Specifically, let the identity b

$$V_{n,t,\phi} \equiv p_{n,t,\phi} \cdot x_{n,t,\phi} \qquad (32)$$

$$\equiv p^*_{n,t} \cdot d_{n,\phi} \cdot b_n \cdot x_n$$

where $p_{n,t,\phi}$, is the unit rental price of the n^{th} capital service of age ϕ at time t, $x_{n,t,\phi}$ is the number of units of capital services provided by the n^{th} capital good of vintage (at time t, $p^*_{n,t}$ is the quality-adjusted, "base" price-index of the n^{th} capital good at tim t, $d_{n,\phi}$, is the deterioration index of the services from good n with vintage ϕ relative to say, age 0 (i.e. it takes the value of unity when the asset is new and declines thereafter) and b_n is the quality index of services from good n at age 0 (defined relative to the service of other new goods) reflecting the effects of embodied technical change. For this reaso b_n is a function of k characteristics $z_{n1},\ldots,z_{nk}$. Also, x_n is the number of standardize units of capital services generated by the n^{th} capital good when it was new (i.e. aged zero) The product $d_{n,\phi} \cdot b_n$ therefore combines the influence of deterioration (the decline i efficiency as capital ages) and embodied technical progress (increasing quality of more re cent vintages).

According to (32), considerable independence exists among the conversion factors $d_{n,($ and b_n. Specifically, deterioration depends on age but not time, and embodied technic change is independent not only of time or age, but also of u and $x_1,\ldots,x_{n-1}$. Hence (3: represents a highly restrictive specification consistent with the simple repackaging hypothes (price independent quality adjustment). Note that under the above assumptions, the pro duct $d_{n,\phi} \cdot b_n$ is a purely technical measure of the relative efficiency or quality of capit services, unaffected by other economic variables.[19] Moreover, in this specification the se vices of old and new capital goods are perfect substitutes up to a factor of proportionalit and under the assumption of cost-minimization the rental prices of alternative capital goo must stand in fixed proportions reflecting their relative marginal product efficiencies (s Robert E. Hall [1971, p.243]).

A related aspect of (32), however, is that the factorization into the two components $d_{n,\phi}$, nd b_n is not unique; this has been shown by Robert E. Hall [1968, 1971]. Essentially, rowth in the product of the two indexes can be identical yet can correspond to differing rowth rates for each of the components; hence an identification problem is present, even n this restrictive simple repackaging (price independent quality adjustment) specification.

One way of eliminating the ambiguity is to adopt a normalization that sets the index f embodied technical change or service quality level equal to well-defined and empirically-ased values for two different vintages at the same time t. For example, if the two models ere identical except for vintage (i.e. $b_n^0 = b_n^1$, $x_n^0 = x_n^1$), then taking ratios of their rental rices in (32) would result in p^*_{nt} dropping out, leaving only the ratio $d^0_{n,\phi}$, to $d^1_{n,\phi}$. Since $_{n,\phi}$, is normalized to unity when $\phi = 0$, taking these ratios would yield well-defined stimates of $d_{n,\phi}$. It is worth noting that for certain assets such as lawn mowers, refrigerators r air conditioners, production runs without model changes often occur for two or more ears; in such cases use of the above procedure would generate clearly identified estimates f the deterioration parameters.

More generally, when different models and varying vintages are compared, Hall [1971] as suggested employing the hedonic technique to account for quality variations using the ocedure described earlier (recall that when two models embody identical characteristics, e of Hall's ratio procedure (21) ensures equal predicted stock quality indexes b_n).

To move from value flows to asset prices, following Muellbauer [1974a] one can substitute 2) into (31) and obtain

$$q_{n,t,\phi} = b_n \cdot p^*_{n,t} \cdot x_n \cdot \sum_{s=0}^{T_n-\phi} \left(\frac{1}{1+r}\right)^s d_{n,\phi+s} \cdot \tag{33}$$

A natural way of defining an index of **depreciation** for the n^{th} capital good (the decline the price of older assets relative to newer ones, observed at the same point in time) is take the ratio of the appropriately discounted expected stream of service values remain-g for the lifetime of the asset to the similarly discounted expected stream of service values

were it new, both evaluated at the same point in time:

$$D_{n,\phi} = \frac{\sum_{s=0}^{T_n-\phi} \left(\frac{1}{1+r}\right)^s d_{n,\phi+s}}{\sum_{s=0}^{T_n} \left(\frac{1}{1+r}\right)^s d_{n,s}} \tag{34}$$

Note that when s = 0, $D_{n,\phi}$ = 1. Multiplying both sides of (34) by the right-sid denominator and substituting into (33) yields

$$q_{n,t,\phi} = p^*_{n,t} \cdot x_n \cdot b_n \sum_{s=0}^{T_n} \left(\frac{1}{1+r}\right)^s d_{n,s}\, D_{n,\phi}\,. \tag{35}$$

Since q represents the value product rather than unit price, now divide both sides of (35 by x_n and denote the resulting unit asset price as $u_{n,t,\phi}$, i.e.

$$u_{n,t,\phi} = \frac{q_{n,t,\phi}}{x_n} = p^*_{n,t} \cdot b_n \sum_{s=0}^{T_n} \left(\frac{1}{1+r}\right)^s \cdot d_{n,s} \cdot D_{n,\phi} \tag{36}$$

$$= p^*_{n,t} \cdot b'_n \cdot D_{n,\phi}$$

where

$$b'_n = b_n \sum_{s=0}^{T_n} \left(\frac{1}{1+r}\right)^s d_{n,s}\,. \tag{37}$$

According to (36), the price of a capital good n of age ϕ at time t is the product of a efficiency-corrected or quality-adjusted rental price index $p^*_{n,t}$ which depends on the tin in which the asset is observed, a depreciation index $D_{n,\phi}$, which varies only with the ag of the asset (since both r and ϕ in (35) are assumed to be constant), and an asset or stoc quality index b'_n that reflects both durability (the discounted time path of deterioratic of the asset) and its quality when new, and which is independent of the year of observatioi

The distinction between the **service** or **flow** quality-adjusted index b_n and the **stock** ʇuality-adjusted index b'_n is important, particularly for the interpretation of intercept terms n hedonic price equations. For example, a slightly different grouping of terms in (36) yields ɩn alternative interpretation.

Specifically, following Hall [1971], regroup (36) as follows:

$$u_{n,t,\phi} = p^{*}_{n,t} \sum_{s=0}^{T_n} \left(\frac{1}{1+r}\right)^s d_{n,s}\, b_n \cdot D_{n,\phi} \tag{38}$$

$$= p^{*\prime}_{n,t} \cdot b_n \cdot D_{n,\phi}$$

lote that b_n appears in (38), while b'_n is in (36). Thus in (38), the first term in brackets $(p'_{n*,t})$ is the efficiency-corrected **stock** price of the new n^{th} asset (rather than the **rental** rice), the b_n term is now the **service** quality (rather than the **stock** quality), and the deprecia-on term $D_{n,\phi}$, remains as before. Muellbauer [1974a, pp.13-14] has argued that if con-ımers are interested in the services yielded by stocks, then over a group of models in x_n ıe services should be perfect substitutes, implying that the rental (rather than asset) prices ıould be in strict fixed proportion to relative service efficiencies. By contrast, to the ex-nt that deterioration time paths and expected lifetimes are different across models, stock ·ices will behave differently from rental prices, and stock prices may not be in fixed pro-ɔrtion to service efficiencies.

The assumption of proportionality of rental prices to service efficiencies is of course ore appealing than the assumption of efficiency-proportionality of stock prices, especially ːıce in utility, production, expenditure or cost functions one is usually interested in ser-ːe quantity flows and prices, rather than stock quantities and stock prices. Note also ıat with the service price specifications, the quality concept of relevance is the stock no-ɩn b'_n (including both durability and quality when new) rather than the Hall's flow con-ːpt b_n. I shall return to this point later.

The hedonic price equations corresponding to (36) and (38) are, since b_n = $(z_{n1}, z_{n2}, \ldots, z_{nk})$,

$$\ln u_{n,t,\phi} = \ln p^{*}_{n,t} + \ln h'_{n}(z_{n1},z_{n2},\ldots,z_{nk}) + \ln D_{n,\phi} \tag{39}$$

and

$$\ln u_{n,t,\phi} = \ln p^{*\prime}_{n,t} + \ln h_{n}(z_{n1},z_{n2},\ldots,z_{nk}) + \ln D_{n,\phi}\ , \tag{40}$$

respectively. Intercept terms in (39) and (40) should be interpreted as quality-adjusted **ser-vice** prices in (39) or quality-adjusted **stock** prices in (40). Note also that in (36), the deteriora-tion term $d_{n,s}$ appears in both the depreciation term $D_{n,\phi}$ and b'_n but not in $p^{*}_{n,t}$, while in (38) the $d_{n,s}$, term appears in $D_{n,\phi}$, and $p^{*\prime}_{n,t}$ but not in b_n. This implies that if deterioration rates $d_{n,s}$ were assumed to differ among alternate types of x_n (say, differen models), consistency would require that in (39) model-specific effects (such as dummy variables) be incorporated in both $\ln D_{n,\phi}$ and $\ln h'_n(z_{n1},\ldots,z_{nk})$ – but not necessarily in the rental price $\ln p^{*}_{n,t}$, while in (40) model effects should be incorporated both in $\ln D_{n,\phi}$ and the stock price $\ln p^{*\prime}_{n,t}$ – but not necessarily in $\ln h_n(z_{n1},\ldots,z_{nk})$.

Finally, it is worth noting that when deterioration is geometric at a constant rate of δ_n the depreciation index $D_{n,\phi}$, also declines geometrically with vintage at the the same rate, i.e

$$D_{n,\phi} = (1 - \delta_n)^{\phi}. \tag{41}$$

(Recall that the depreciation index $D_{n,\phi}$, compares retained value proportions of asse identical in all respects except vintage at a given point in time, and not the decline in th value of the asset as it ages between two different points in time; this implies that the di ference between $D_{n,\phi}$, and $D_{n,\phi+s}$ depends only on δ_n, and not on r.) Inserting (41) int (39) then yields the estimable hedonic price equation

$$\ln u_{n,t,\phi} = \ln p^{*}_{n,t} + \ln h'_{n}(z_{n1},\ldots,z_{nk}) + \ln (1 - \delta_n) \cdot \phi\ , \tag{42}$$

an equation relating used asset price to characteristics and age. After adding an independen ly and identically normally distributed random disturbance term to an equation like (42 Muellbauer [1971a, 1974a] has estimated parameters employing data on prices of us

capital goods (farm tractors) observed at different times, plus dummy variables for vintages, models, and time; Hall [1971] added to (42) physical characteristics of Ford and Chevrolet pickup truck models. Tests for the validity of the simple repackaging hypothesis were conducted by Muellbauer by testing whether interaction terms (e.g., model-time, depreciation-time) had estimated coefficients significantly different from zero.

It is worth noting here that the above analysis of durable good quality is based on the simple repackaging (price independent) quality aggregation hypothesis. Hence this framework would not be appropriate for analysis of interesting and important issues such as the determination of whether and to what extent fuel price increases have altered the economic depreciation patterns of various energy-using assets such as autos, refrigerators, or air conditioner models since 1970. To undertake such an analysis would require relaxing the simple repackaging hypothesis (price independent quality adjustment), and then allowing the conversion ratio b_n' in (36) to depend on prices of other commodities such as gasoline or electricity. Moreover, and this could be very important empirically, since b_n' embodies a stock notion rather than a flow concept (see (37)), it would be necessary to specify that the energy cost variable in the conversion function b_n' reflect discounted **lifetime** (rather than remaining **annual**) fuel costs were it new.

Let us now briefly consider extension of this durable good framework to the more general variable repackaging (price independent quality adjustment) hypothesis type of depreciation, using the example of autos and fuel prices. Following the earlier analysis, denote the quantity of, say, gasoline fuel as x_{n-1}, and its price as p_{n-1}. Under the simple repackaging (price independent) hypothesis in (37),

$$b_n' = h_n'(z_{n1}, z_{n2}, \ldots, z_{nk}) \sum_{s=0}^{T_n} \left(\frac{1}{1+r}\right)^s d_{n,s} \qquad (43)$$

where the $z_{n1}, z_{n2}, \ldots, z_{nk}$ are independent of u and commodity quantities $x_1, x_2, \ldots, x_{n-1}$ or commodity prices $p_1, p_2, \ldots, p_{n-1}$. Note that each of the characteristics in (43) is implicitly assumed to generate services that deteriorate over time at the same rate $d_{n,s}$ (although of course $d_{n,s}$ is permitted to vary with s unless constant geometric deterioration is assumed).

One empirically tractable generalization of (43) consistent with the variable repackagin (price dependent) hypothesis discussed earlier (see equations (22)-(30) above) is to specif that the b_n conversion or quality aggregation function depends not only on $z_{n1}, z_{n2}, \ldots, z_{nk}$ but also on x_{n-1} (or, equivalently, p_{n-1}). In such a case (43) becomes

$$b_n' = h_n'(z_{n1}, z_{n2}, \ldots, z_{nk}, p_{n-1}) \sum_{s=0}^{T_n} \left(\frac{1}{1+r}\right)^s d_{n,s} \qquad (44)$$

and the hedonic price equation (39) becomes

$$\ln u_{n,t,\phi} = \ln p_{n,t}^* + \ln h_n'(z_{n1}, z_{n2}, \ldots, z_{nk}, p_{n-1}) + \ln D_{n,\phi} \qquad (45)$$

While it is again implicitly assumed in (44) that the adverse effects of fuel price increase deteriorate over vintages at the same rate as other characteristics, an additional featur of (44) and (45) is that the numerical values of p_{n-1} will vary over time for given model unlike other engineering characteristics; hence in the variable repackaging input qualit case b_n' is no longer necessarily constant over time. This is attractive, for it permits qual ty adjustment between ''gas guzzlers'' and ''gas misers'' to vary with the price of gasolin

It is worthwhile noting, incidentally, that hedonic equations similar to (45) have recen ly been estimated using second-hand automobile market data by, among others, Jam Kahn [1982], George Daly and Thomas Mayor [1983] and Zvi Griliches and Makota Oh [1983]. Their regression results suggest quite clearly that the more general price-depende (variable repackaging) specification (45) is preferable to that of (39), for not only automobile prices depend on engineering design and performance characteristics, but the also depend on the price of gasoline.

IV. On the Interpretation of Coefficients in Hedonic Price Equations

In the previous paragraphs I have related quality adjustment for durable and non-durab goods in demand analysis to the well-known hedonic price literature. I now briefly digre to consider conditions under which parameters from hedonic price equations can be inte preted unambiguously as reflecting demand (rather than cost or supply) conditions.

Suppose that for a particular durable or non-durable commodity there existed K detail-
ed engineering, design, performance, or other "quality" characteristics. Denote measures of these K attributes as $z_1, z_2, \ldots, z_K$. Let each model n of vintage v embody a particular configuration of these characteristics. In the hedonic formulation the price of a durable good, u_{nv}, is decomposed into implicit (shadow) prices (denoted $c_1, c_2, \ldots, c_K$) corresponding with the quantity measures $z_1, z_2, \ldots, z_K$ of the attributes, i.e.

$$u_{nv} = f(c_1, z_1, c_2, z_2, \ldots, c_K, z_K) \,. \qquad (46)$$

Recall that under the variable repackaging (price dependent) quality aggregation hypothesis, the list of characteristics in (46) might include quantities (or prices) of commodities related to the engineering characteristics, e.g., fuel prices.

In order empirically to link hedonic price analysis with the modern flexible functional form demand analysis, in principle it is important that coefficients of the hedonic price equations (45) and (46) be properly interpreted as representing demand function parameters. In practice, problems of interpretation arise because in general both supply and demand functions exist for the good/characteristic combinations. Since the hedonic equation (45) or (46) is essentially a reduced form, the existence of varying imperfect market structures may make it impossible in general to retrieve unique structural estimates of demand or supply function parameters using hedonic regression equations based on observed market price, sales and characteristic data.[20]

If the supplying market were composed of identical and perfectly competitive firms and the production of attributes were characterized by constant returns to scale, then the parameters of (46) could be interpreted as representing the average and marginal costs of characteristics. In such cases prices would of course be supply-determined. As Sherwin Rosen [1974] has noted, however, product markets for durable goods are likely to involve non-identical firms selling slightly differentiated new products; others have noted that differentiated markets for durable goods often tend to be oligopolistic in nature.[21] Moreover, for successful new product innovations embodying a novel configuration of characteristic combinations, temporary monopoly profits may exist as rewards to innovation, thereby driving a wedge between marginal costs of production and market price.

On the other hand, if the supply curves of the slightly differentiated products or models (each embodying alternative combinations of characteristics) were perfectly inelastic, then the market demand and supply curves would intersect at different levels for each model (characteristic combination). In such a case the structure of prices would be demand-determined, and the difference in price levels among models could be interpreted unambiguously as providing implicit measures of consumers' evaluations of the characteristic combinations, i.e. as well-identified estimates of demand function parameters.

When, however, supply is neither perfectly elastic nor perfectly inelastic, prices are jointly determined by supply and demand. In such cases special care and additional assumptions must be made in order to extract from reduced form hedonic price equations identifiable parameters of the underlying cost and demand functions. The most obvious alternative approach is to estimate jointly structural supply and demand functions, where the supply function is based on a multi-attribute or multi-product cost function and the demand functions also incorporate these characteristics. Often, however, the required data are not available.

The identification issue in a reduced form hedonic equation was addressed in an important paper by Sherwin Rosen [1974], wherein he proposed a two-step instrumental variable procedure. Recently James N. Brown and Harvey S. Rosen [1982] have qualified some of Rosen's results, suggesting that identification of cost and demand function parameters for new products is not always possible with Rosen's two-step instrumental variable estimator.

While all these authors deal extensively with interpretation of hedonic regression parameters based on new product data, none appear explicitly to have considered the possibility of incorporating into the analysis the fact that second-hand, leasing or rental markets provide additional economic information that can facilitate identification of structural demand or cost function parameters.

Used or secondary markets across space or time are of considerable relevance, since supply is almost perfectly inelastic. Once a production run of a particular new car, truck, tractor, or other equipment model is made and sold, durability of the equipment implies that unless it is scrapped, its total quantity is fixed. Each year the owner can be envisaged as making a choice between renting the asset to himself or renting it to someone else. To the extent that scrapping is not empirically significant (which empirically is the case for autos up to about eight years and for farm tractors up to about 12 years), empirical analysis of used asset markets provides reasonably reliable estimates of demand function parameters, for supply is essentially perfectly inelastic.[22]

One other cautionary note in this context involves allowance for inter-actions between new and used markets. Often a particular piece of equipment in the used market is considered in isolation from the new market. In such a case an outward shift in the demand curve for, say, used fuel-efficient models is viewed as having no immediate effect in the new market, in spite of the price signal generated by rising relative prices of fuel-efficient used models. If, however, new and used models were at least partial substitutes and if the supply of new fuel-efficient models were rapidly responsive to relative price signals generated in the used market, analysis confined to the used market would no longer contain information only on demand, i.e. used model prices would again be determined jointly by supply and demand. However, such jointness would require rapid responses by durable goods manufacturers, which is somewhat unlikely due to the long lead times often required to introduce new models.

It is clear, therefore, that market structure affects the interpretation of hedonic price equations in a very important manner. Identification of cost or demand function parameters may be difficult even when Rosen's two-step instrumental variable estimator is employed. However, identification of demand parameters can be facilitated when data across space or time on used or second-hand markets are exploited, since in those cases supply may be inelastic and prices will reflect only demand parameters. Note also that if data on used markets are available at different points in time (say, a pooled cross-section, time series data set providing the history of used prices for various models), one could employ the hedonic technique to test whether consumers' preferences and evaluations have changed over time.

A final issue in interpretation of hedonic price equation coefficients concerns the choic of functional form. As noted in the Introduction, it is useful to view hedonic regression as generating a "quality-adjusted" price index for durable or non-durable goods, whic implies that the theoretical foundations of the hedonic technique should be closely relate to the economic theory of index numbers and the "true cost-of-living indexes". Indeed the hedonic equations can be viewed as aggregating component characteristics and price into an aggregate scalar index of quality. In turn, since the theory of index numbers i closely intertwined with the theory of cost, production and utility,[23] it follows tha economic theory might imply certain restrictions on the functional form of the hedoni regression equation.

In a series of papers, John Muellbauer [1971a,b; 1972; 1974a,b; 1975a,b] has show that in fact economic theory does place testable parametric restrictions on the functiona form of hedonic regression equations when such equations are interpreted as providin input quality-adjusted price indexes. For example, in Muellbauer [1974b] it is shown tha a logical contradiction occurs when one assumes a semi-logarithmic relationship betwee prices and characteristics and then also allows the parameters in the relationship to var from year to year.[24] Also, the hedonic price equation should be homogeneous of degre one in prices of its components. Another problem with the semi-logarithmic form is tha with it the identity between value and the multiplicative product of prices and quantitie may not be globally preserved. Note, however, that in general it is not required that th hedonic price equation be homogeneous of degree one in the quantities of its components Thus on the basis of economic theory, either linear-linear, linear-quadratic, log-log linear or log-log quadratic functional forms are preferable to the semi-logarithmic representa tion of log price on a linear function of the characteristics, although choice among th set of preferable forms on the basis of theory is not yet clear.

In the previous paragraphs I have digressed briefly to review recent literature on th interpretation of coefficients in hedonic regression equations, and have emphasized th role of economic theory and second-hand markets in facilitating identification of deman function parameters. I now proceed to illustrate a number of ways in which the hedoni technique can be incorporated into modern flexible functional form empirical deman

analysis and price index construction, providing both richness in characteristic detail yet parsimony in parameterization.

V. Towards Empirical Implementation

Earlier I developed an intuition as to what precisely is meant by the term "quality", and how quality aspects for non-durable and durable goods relate to the modern theory of commodity or input demand. In this section I turn to outlining possibilities for implementing empirical research on quality-quantity demand models.

At the outset, it is useful to emphasize again the structural framework that has been developed concerning interpretation of hedonic price equations. Specifically, in this paper I have been concerned primarily with the interpretation of an hedonic equation within the theory of demand; supply and general equilibrium aspects have not been addressed in a detailed manner. The specification of an hedonic price equation has been shown to be equivalent to the specification of quality conversion functions for commodities or inputs. In turn, these quality conversion functions have been specified to be either price-independent (corresponding to the simple repackaging hypothesis) or price-dependent (the variable repackaging hypothesis).

In the case of price-independent quality conversion, the implied hedonic price equation is of the familiar form of quality unadjusted price as a function of characteristics and attributes; the intercept term in such an equation represents price per standardized or quality-adjusted unit. Note that such a structural equation is of the same form as the numerous "reduced form" hedonic equations surveyed by, for example, Zvi Griliches [1971a,b].

By contrast, when quality conversion is price-dependent, the implied hedonic price equation consistent with this theory of demand relates quality unadjusted price not only to characteristics or attributes, but also to the price (or quantity) of another commodity or input. Hence price is a regressor in this structural hedonic demand equation. Griliches [1971a, p.5] has expressed considerable reservations about having market-determined prices or quantities as regressors in an hedonic price equation, but his vantage is clearly one of reduced form rather than structural analysis.

It is worth noting once again that within the last two years a number of hedonic studie have appeared in the empirical literature with prices as regressors; see, for example, th used auto studies by Kahn [1982], Daly-Mayor [1983] and Griliches-Ohta [1983]. One im portant empirical implication of this paper is that such structural hedonic equations hav a clear and interesting interpretation, for in effect they provide parameter estimates o price-dependent quality conversion equations consistent with the theory of demand.

A second important empirical implication emerging from the previous sections concern the interpretation of intercept terms in hedonic equations for durable goods. As has bee noted earlier by Muellbauer [1974a], if one believes that durable good services rather tha durable good stocks are perfect substitutes, then rental rather than asset prices should b proportional to service efficiencies. This implies both that the relevant quality concept i the stock notion $b_n^{\prime}$ (see (37)) rather than flow concept b_n and that the intercept term refer to the quality-adjusted rental rather than asset price.

The discussion to this point has concerned itself primarily with the interpretation o parameters in structural single-equation hedonic price equations. A more significan empirical implication of the approach presented above, however, concerns the joint effi cient estimation of structural demand parameters and quality conversion coefficients i systems of demand equations with testable cross-equation parametric constraints. To se this, assume the utility function is of the form where only the n^{th} commodity is qualit adjusted, i.e.

$$u = F(x_1, x_2, \ldots, x_{n-1}, x_n^*) \,. \tag{47}$$

and where the budget constraint is $y = \sum_{i=1}^{n-1} p_i \cdot x_i + p_n^* \cdot x_n^*$.

Define the indirect utility function as

$$v = G(p_1, p_2, \ldots, p_{n-1}, p_n^*, y) \,, \tag{48}$$

where v is the maximum attainable level of utility given the budget constraint y and input prices $p_1, p_2, \ldots, p_n^*$. Denote the normalized prices as P,

$$P = [p_1/y, p_2/y, \ldots, p_{n-1}/y, p_n^*/y] \tag{49}$$

$$= [P_1, P_2, \ldots, P_{n-1}, P_n^*] \,.$$

Now let the indirect utility function (48) be of the translog form,[25]

$$\ln v = \alpha_0 + \sum_{i=1}^{n^*} \alpha_i \ln P_i + \frac{1}{2} \sum_{j=1}^{n^*} \beta_{ij} \ln P_i \ln P_j \tag{50}$$

where $\beta_{ij} = \beta_{ij}$.

Now specify the quality-adjusted price P_n^* as P_n/b_n, where $P_n = p_n/y$ and $b_n = h_n(z_n)$. Initially, assume that the vector $z_n = [z_{n1}, a_{n2}, \ldots, z_{nk}]$ contains only characteristics, and no prices or quantities of other commodities; this is consistent with price-independent quality adjustment (in Muellbauer's terminology, the simple repackaging hypothesis). Moreover, in order to be compatible with the logarithmic translog form, next specify that the hedonic price equation be of the log-log form,

$$\ln P_n = \ln P_n^* + \sum_{k=1}^{K} b_{nk} \ln z_{nk} \tag{51}$$

which of course implies

$$\ln P_n^* = \ln P_n - \sum_{k=1}^{K} b_{nk} \ln z_{nk} \tag{52}$$

Now subsitute (52) into (50), and then use Roy's [1943] identity in logarithmic form,

$$\frac{p_i x_i}{y} = \frac{-\partial \ln v}{\partial \ln P_i} \cdot \frac{\partial \ln v}{\partial \ln y}, \quad i = 1, \ldots n \tag{53}$$

to obtain the optimal budget shares which, after substitution of (52) yields

$$\frac{p_i x_i}{y} = \frac{\alpha_i + \sum_{j=1}^{n^*} \beta_{ij} \ln P_j}{\sum_{j=1}^{n^*} \alpha_j + \sum_{j=1}^{n^*} \sum_{i=1}^{n^*} \beta_{ij} \ln P_i \ln P_j}, \quad i=1,\ldots n \tag{54}$$

In order that the budget share equations (54) be homogenous of degree zero in the parameters, I adopt the normalization that

$$\sum_{i=1}^{n^*} \alpha_i = -1.$$

Note that when (52) is substituted back into (54), the budget share equations depend not only on the normalized prices P_i, $i=1,\ldots,n$, but also on the characteristics z_{nk}, $k=1,\ldots,K$; moreover, there are testable cross-equation restrictions on the hedonic parameters b_{nk}, which appear in each of the share equations. Hence when the structural hedonic price (quality-quantity adjustment) framework is integrated with the modern theory of demand, characteristics enter the system of budget share equations with testable cross-equation restraints.

Although these cross-equation constraints are present in rather general formulations they also occur under more restrictive conditions. Consider, for example, the case wher homotheticity (unitary income or expenditure elasticities) is imposed on the translog indirect utility function; this implies the parametric restrictions

$$\sum_{j=1}^{n^*} \beta_{ij} = 0, \quad i=1,2,\ldots,n^* \tag{55}$$

When these homotheticity restrictions are substituted into the budget share equations (54) one obtains the simpler system,

$$\frac{p_i x_i}{y} = -\alpha_i - \sum_{j=1}^{n-1} \beta_{ij} \ln P_j - \beta_{in*} (\ln P_n - \sum_{k=1}^{K} b_{nk} \ln z_{nk}), \quad i=1,\ldots,n. \tag{56}$$

which makes more clear the presence of characteristics and b_{nk} in each of the share equation, i.e. the existence of testable cross-equation parameter restrictions. Note that when used with, for example, time series data on P_i and z_n, econometric estimation occurs for both the structural (β_{ij}, α_i) and hedonic (b_{nk}) parameters.[26] This demonstrate that modern flexible form demand analysis can be integrated with hedonic price analysis in an empirically implementable form with testable cross-equation parameter restrictions.

Suppose, for example, that the n^{th} commodity were food, and that the z_n vector consisted of a set of nutritional variables such as fat, protein, vitamin, sodium and caloric content. In such a case, these food nutritional variables would appear in each of the estimable equations with cross-equation constraints. The null hypothesis that "quality" (nutritional content) does not matter would correspond with the joint null hypothesis that $b_{nk} = 0$, $k = 1,\ldots,k$. Hence such an equation system would reflect two basic premises: (i) if quality is important, it should be evident in quantity or share equations; and (ii) economic theory imposes testable parametric restrictions on the way in which quality enters these quantity or share equations.

The above example of empirical implementation of the quality-quantity demand framework was based on the assumption that the quality adjustment function was price-independent. I now briefly outline generalization to price-dependent quality adjustment.

Suppose, for example, that the n^{th} commodity in the utility function (47) referred to the net services of air conditioners. However, since the net services obtained from a durable good such as air conditioners depend on operating costs such as the costs of electricity, it is reasonable to specify that conversion ratios among air conditioners having differing energy-efficiency ratios (EER's) depend on the price of electricity, and thus that the rental price of air conditioners in the corresponding hedonic price equation be a function both of the characteristic EER and the price of electricity. This corresponds to the case of price-dependent quality adjustment.

Given data on the distribution and levels of air conditioners with differing EER's, assumptions concerning the constant geometric rate of deterioraton δ, the discount rate r and the price of electricity P_{Elec}, one could use (44) to specify a present-valued operating cost variable for air conditioners as

$$OC = g(P_{Elec}, r, \delta, EER) \tag{57}$$

and then specify an hedonic equation of the form

$$\ln P_n = \ln P_n^* + b_{n0} \ln OC. \tag{58}$$

Solving for $\ln P_n^*$,

$$\ln P_n^* = \ln P_n - b_{n0} \ln OC, \tag{59}$$

one could substitute back into the indirect utility function (50), employ Roy's identity, and then obtain budget share equations for electricity, the services of air conditioners, and all other commodities, each as a function of total expenditure, commodity prices and OC. Again, the parameter b_{n0} would appear in each of the share equations implying testable cross-equation constraints; moreover, whether quality mattered could be tested simply as whether b_{n0} was statistically different from zero.

The above examples illustrate the empirical research potential made possible by the integration of modern demand analysis with hedonic price analysis. This integration also has clear implications for index number construction, provided of course that the resulting index number be interpreted within the context of economic "true" cost-of-living indexes.[27] As an example, one could incorporate into the price index of meat studies by Christensen-Manser [1976, 1977] a number of nutritional variables; the resulting conditional price indexes for meat (holding u fixed) would then depend explicitly on structural substitution parameters of demand for meat and on the hedonic coefficients of the nutritional variables.

VI. Concluding Remarks

It has been the purpose of this paper to present and discuss a theoretical framework through which durable and non-durable commodities can be quality-adjusted through the integration of hedonic price analysis with modern flexible functional form demand analysis,

and thereby to relate quality adjustment to the economic theory of index numbers.

The examples presented in this paper have been drawn primarily from the theory of consumer demand. As was noted in Section II, however, this framework is easily transferable to the analysis of producer costs and production. Potential empirical applications of this framework to the factor demand, productivity, and multiple output context have been outlined in Section V of Berndt [1983a]; Berndt [1983b] provides empirical implementation based on the price-independent quality adjustment hypothesis for U.S. manufacturing, 1958-77. It might also be noted that classic empirical studies of production behavior in the U.S. can now be re-interpreted within the integrated hedonic-structural demand approach of this paper; see, for example, Griliches [1970] on the quality of labor as a function of educational attainment.

A number of analytical extensions are also suggested by this research. For example, although this paper has employed the assumption of static optimization, recent work on dynamic factor demand models[28] suggests that generalization to dynamic optimization is feasible and empirically implementable. Specifications of expectations formation, however, will naturally affect the way in which capital quality, quantity and rental price should be measured. Research on this topic is clearly important.

Another area for fruitful research concerns aggregation over consumers rather than commodities. Specifically, in much recent consumer budget research, individual family units of varying demographic composition have been re-weighted using family equivalence scales; see, for example, Angus Deaton and John Muellbauer [1980]. The relationship between family equivalence scaling and quality adjustment is not yet clear, and deserves careful attention. If these two notions could be combined, it might be possible to generate quality-adjusted price indexes for various demographic groups as a function of the distribution of expenditures, characteristics, and demographic variables.

Finally, with respect to recent developments in the economic theory of index numbers (see, for example, Diewert [1976, 1981], Pollak [1983] and Triplett [1983]), the framework adopted here involves aggregation of characteristics into a scalar quality measure and thus

places separability-type restrictions on the structure of utility functions. These separability restrictions need to be examined more carefully, along with their implications for the construction of index numbers. For example, the price-independent quality adjustment specification could be viewed as placing greater separability restrictions on the functional structure than does price-dependent quality adjustment.[29]

Issues of quality adjustment via hedonic price analysis have a long and distinguished history in the literature on index number construction. In recent years the modern theory of consumer demand has been linked with the economic theory of index numbers. In this paper I have attempted to contribute to both these areas by integrating hedonic price analysis with modern flexible functional form demand analysis. Since the resulting specifications incorporate characteristic data yet still remain relatively parsimonious in parameterization, the potential for new empirical research based on this integration is rich and exciting.

Footnotes

[1] S.L. Horner [1939, p.5].

[2] A.T. Court [1939, p.116].

[3] *Ibid.*, Footnote 3, p.101.

[4] *Ibid.*, pp.101-103.

[5] *Ibid.*, p.107.

[6] *Ibid.*

[7] *Ibid.*, pp.101-103, 112. It is not always the case, however, that quality adjustment reduces the rate of growth of the price index; see M.L. Burstein [1961] and Jack Triplett [1971a,b].

[8] See, for example, S.M. Du Brul [1939, pp.126-130].

[9] For an earlier attempt at quality adjustment using regression techniques, see the study on vegetable prices and quality by F.V. Waugh [1929].

[10] See, for example, W.M. Gorman [1956] and Richard Stone [1956].

[11] See W. Erwin Diewert [1971], Laurits R. Christensen, Dale W. Jorgenson and Lawrence J. Lau [1971], and Ernst R. Berndt and Mohammed S. Khaled [1979]. For a history and brief survey of earlier contributions, see Barry C. Field and Ernst R. Berndt [1981].

[12] John Muellbauer [1975b, p.282]. For a theoretical attempt to "rationalize" hedonic equations in the context of new goods, see W. Erwin Diewert [1980, pp.503-505]; also, on the production side, see Makota Ohta [1975].

[13] For a discussion of such specification issues, see Melvyn Fuss, Daniel McFadden and Yair Mundlak [1978].

[14] See Elizabeth E. Bailey and Ann F. Friedlaender [1982] for a brief survey of econometric studies estimating economies of scale and economies of scope in multi-product firms, including quality adjustment. Also see Richard H. Spady [1979], Ann F. Friedlaender and Richard H. Spady [1981], J.S. Wang Chiang [1981], and J.S. Wang Chiang and Ann F. Friedlaender [1982]. In the context of telecommunications, see Michael Denny *et al.* [1981a,b].

[15] Diewert's presentation does not introduce b explicitly, but these dual relationships are compatible with it. See McFadden [1978].

[16] For a discussion of the simple repackaging hypothesis in the context of n (rather than just one) commodities, see John Muellbauer [1974a; 1975a].

[17] This treatment of hedonics within an explicit theory of production provides an effective counter-example to the concerns of including market-determined quantities in an hedonic price equation voiced by, in particular, Zvi Griliches [1971a, p.5].

[18] In yet a different version of the variable repackaging hypothesis, Muellbauer specifies b_n to be independent of y and $x_1, x_2, \ldots, x_{n-1}$, but dependent on x_n. Under constant returns to scale, however, in this case the simple and variable repackaging hypotheses coincide; see Muellbauer [1975a, footnote 6, p.42].

[19] For a discussion of these assumptions, see Dale W. Jorgenson [1974] and Martin S. Feldstein and Michael Rothschild [1974]. Note also that it would be relatively simple to add disembodied technical change to the above specification; see Hall [1968].

[20] The importance of market structure in identifying supply or demand parameters was emphasized already in 1961 by Meyer L. Burstein, discussed briefly by Irma Adelman and Zvi Griliches [1961], yet received very little empirical or theoretical attention until Sherwin Rosen [1974].

[21] For an empirical example of hedonic cost function estimation and identification in imperfect markets under a constant mark-up assumption, see Makota Ohta [1975]; also see Makota Ohta and Zvi Griliches [1975].

[22] For empirical hedonic studies of used markets under the assumption of inelastic supply, see Phillip Cagan [1965], Robert E. Hall [1971], Charles R. Hulten and Frank C. Wykoff [1981a,b], John Muellbauer [1971a], Makota Ohta and Zvi Griliches [1975], James Kahn [1982], George Daly and Thomas Mayor [1983] and Zvi Griliches and Makota Ohta [1983].

[23] See W. Erwin Diewert [1976, 1980, 1981], Robert A. Pollak [1983], Robert E.B. Lucas [1975] and Jack E. Triplett [1976].

[24] This is a very common practice. See, for example, the studies surveyed in Zvi Griliches [1971a,b].

[25] Other flexible forms are of course available. For an empirical comparison, see Berndt, Darrough and Diewert [1977] and Berndt-Khaled [1979].

[26] It would also be possible, of course, to obtain estimates of the hedonic parameters from a different body of data, substitute these into (56), and then estimate only the structural parameters α_i and β_{ij} in (56); the alternative suggested here within a system of equations has the advantage of permitting more efficient estimation.

[27] For a survey of the economic theory of index numbers, see W. Erwin Diewert [1981].

[28] This literature is surveyed in Ernst R. Berndt, Catherine J. Morrison, and G. Campbell Watkin- [1981]; more recent contributions include Catherine J. Morrison [1982] and Robert S. Pindyck and Julio J. Rotemberg [1982].

[29] Under price-independent quality adjustment, the functional structure is inherently asymmetric and has been called weakly recursive separability by George Lady and David Nissen [1968]; also see Charles Blackorby, Daniel Primont and R. Robert Russell [1975] and the discussion of group wise separability by Dale W. Jorgenson and Lawrence J. Lau [1975].

References

Adelman, Irma and Zvi Griliches [1961], "On an Index of Quality Change", *Journal of the American Statistical Association*, Vol. 56, pp.535-548.

Bailey, Elizabeth E. and Ann F. Friedlaender [1982], "Market Structure and Multiproduct Industries", *Journal of Economic Literature*, Vol. 20, No. 3, September, pp.1024-1048.

Berndt, Ernst R. [1983a], "Quality Adjustment in Empirical Demand Analysis", Cambridge, MA: M.I.T. Sloan School of Management, Working Paper No. 1397-83, January.

———— [1983b], "Electrification, Energy Quality, and Productivity Growth in U.S. Manufacturing", Cambridge, MA: M.I.T. Sloan School of Management, Working Paper No. 1421-83, March.

———— W. Erwin Diewert and Masoko N. Darrough [1977], "Flexible Functional Forms and Expenditure Distributions: An Application to Canadian Consumer Demand Functions", *International Economic Review*, Vol. 18, No. 3, October, pp.651-675.

———— and Mohammed S. Khaled [1979], "Parametric Productivity Measurement and Choice Among Flexible Functional Forms", *Journal of Political Economy*, Vol. 87, No. 6, December, pp.1220-1245.

———— Catherine J. Morrison and G. Campbell Watkins [1981], "Dynamic Models of Energy Demand: An Assessment and Comparison", Chapter 12 in Ernst R. Berndt and Barry C. Field, eds., *Modeling and Measuring Natural Resource Substitution*, Cambridge: The M.I.T. Press, pp.259-289.

Blackorby, Charles, Daniel Primont and R. Robert Russell [1975], "Budgeting, Decentralization and Aggregation", *Annals of Economic and Social Measurement*, Vol. 4, No. 1, pp.23-44.

Brown, James N. and Harvey S. Rosen [1982], ''On the Estimation of Structural Hedonic Price Models'', *Econometrica*, Vol. 50, No. 3, May, pp.765-768.

Burstein, M.L. [1961], ''Measurement of Quality Change in Consumer Durables'', *Manchester School of Economics and Social Studies*, Vol. 29, pp.267-279.

Cagan, Phillip [1965], ''Measuring Quality Changes and the Purchasing Power of Money: An Exploratory Study of Automobiles'', *National Banking Review*, Vol. 3, pp.217-236. Reprinted in Zvi Griliches, ed., *Price Indexes and Quality Change*, Cambridge: Harvard University Press, 1971, pp.215-239.

Christensen, Laurits R., Dale W. Jorgenson and Lawrence J. Lau [1971], ''Conjugate Duality and the Transcendental Logarithmic Production Function'' (abstract), *Econometrica*, Vol. 39, No. 4, July, pp.255-256.

__________ and Marilyn Manser [1976], ''Cost-of-Living Indexes and Price Indexes for U.S. Meat and Produce, 1947-71'', in Nester E. Terleckyj, ed., *Household Production and Consumption*, Studies in Income and Wealth No. 40, New York: National Bureau of Economic Research, pp.399-446.

__________ [1977], ''Estimating U.S. Consumer Preferences for Meat with a Flexible Utility Function'', *Journal of Econometrics*, Vol. 5, No. 1, January, pp.37-54.

Court, A.T. [1939], ''Hedonic Price Indexes with Automobile Examples'', in *The Dynamics of Automobile Demand*, New York City: General Motors Corporation, pp.95-117.

Daly, George C. and Thomas H. Mayor [1983], ''Reason and Rationality during Energy Crises'', *Journal of Political Enonomy*, Vol. 91, No. 1, February, pp.168-181.

Deaton, Angus and John Muellbauer [1980], *Economics and Consumer Behavior*, Cambridge: Cambridge University Press.

Denny, Michael, Carol Everson, Melvyn Fuss and Leonard Waverman [1981], "Estimating the Effects of Diffusion of Technological Innovations in Telecommunications: The Production Structure of Bell Canada", *Canadian Journal of Economics*, Vol. 14, No. 1, February, pp.24-43.

———— Melvyn Fuss and Leonard Waverman [1981], "The Measurement and Interpretation of Total Factor Productivity in Regulated Industries, with an Application to Canadian Telecommunications", Ch. 8 in Thomas Cowing and Rodney Stevenson, eds., *Productivity Measurement in Regulated Industries*, New York: Academic Press, pp. 179-218.

Diwert, W. Erwin [1971], "An Application of the Shephard Duality Theorem: A Generalized Leontief Production Function", *Journal of Political Economy*, Vol. 79, May/June, pp.481-507.

———— [1976], "Exact and Superlative Index Numbers", *Journal of Econometrics*, Vol. 4, No. 2, May, pp.115-145.

———— [1980], "Aggregation Problems in the Measurement of Capital", in Dan Usher, ed., *The Measurement of Capital*, Chicago: The University of Chicago Press, Ch. 8, pp.433-528.

———— [1978], "Duality Approaches to Microeconomic Theory", Discussion Paper 78-09, University of British Columbia, Department of Economics, March. Forthcoming in Kenneth J. Arrow and Michael D. Intriligator, eds., *Handbook of Mathematical Economics*, Amsterdam: North-Holland Publishing Company.

———— [1981], "The Economic Theory of Index Numbers: A Survey", Ch. 7 in Angus Deaton, ed., *Essays in the Theory and Measurement of Consumer Behavior in Honour of Sir Richard Stone*, Cambridge: Cambridge University Press, pp.163-208.

u Brul, S.M. [1939], "Statement of the Problem", in *The Dynamics of Automobile Demand*, New York City: General Motors Corporation, pp.3-20.

Feldstein, Martin S. and Michael Rothschild [1974], ''Towards an Economic Theory of Replacement Investment'', *Econometrica*, Vol. 42, No. 3, May, pp.393-423.

Field, Barry C. and Ernst R. Berndt [1981], ''An Introductory Review of Research on the Economics of Natural Resource Substitution'', Ch. 1 in Ernst R. Berndt and Barry C. Field, eds., *Modeling and Measuring Natural Resource Substitution*, Cambridge, MA: The M.I.T. Press, pp.1-14.

Fisher, Franklin M. and Karl Shell [1968], ''Taste and Quality Change in the Pure Theory of the True Cost-of-Living Index'', in J.N. Wolfe, ed., *Value, Capital, and Growth: Essays in Honour of Sir John Hicks, Edinburgh:* University of Edinburgh Press. Reprinted as Ch. 2 in Zvi Griliches, ed., *Price Indexes and Quality Change*, Cambridge: Harvard University Press, 1971, pp.16-54.

Friedlaender, Ann F. and Richard H. Spady [1981], *Freight Transport Regulation: Equity, Efficiency and Competition in the Rail and Trucking Industries*, Cambridge, MA: The M.I.T. Press.

Fuss, Melvyn, Daniel McFadden and Yair Mundlak [1978], ''A Survey of Functional Forms in the Economic Analysis of Production'', Ch. II.1 in Melvyn Fuss and Daniel McFadden, eds., *Production Economics: A Dual Approach to Theory and Applications*. Amsterdam: North-Holland Publishing Company, Vol. 1, pp.217-268.

Gorman, W.M. [1956], ''A Possible Procedure for Analysing Quality Differentials in the Egg Market'', Discussion Paper No. B4, London School of Economics (a re-issue in January 1976 of the 1956 paper).

Griliches, Zvi [1970], ''Notes on the Role of Education in Production Functions and Growth Accounting'', in W. Lee Hansen, ed., *Education, Income, and Human Capital*, Studies in Income and Wealth No. 35, New York: Columbia.

__________ [1971a], ''Introduction: Hedonic Prices Revisited'', Ch. 1 in Zvi Griliches ed., *Price Indexes and Quality Change*, Cambridge: Harvard University Press, pp.3-15

———— [1971b], ''Hedonic Price Indexes for Automobiles: An Econometric Analysis of Quality Change'', Ch. 3 in Zvi Griliches, ed., *Price Indexes and Quality Change*, Cambridge: Harvard University Press, pp.55-87.

———— and Makota Ohta [1983], ''Gasoline Price Changes and Consumers' Relative Quality Evaluations'', Mimeo, National Bureau of Economic Research, May.

Hall, Robert E. [1968], ''Technical Change and Capital from the Point of View of the Dual'', *Review of Economic Studies*, Vol. 35, No. 1, January, pp.35-46.

———— [1971], ''The Measurement of Quality Change from Vintage Price Data'', Ch. 8 in Zvi Griliches, ed., *Price Indexes and Quality Change*, Cambridge: Harvard University Press, 1971, pp.240-271.

Hanoch, Giora [1978], ''Polar Functions with Constant Two Factors-One Price Elasticities'', Ch. II.3 in Melvyn Fuss and Daniel McFadden, eds., *Production Economics: A Dual Approach to Theory and Applications*, Amsterdam: North-Holland Publishing Company, Vol. 1, pp.287-310.

Horner, S.L. [1939], ''Significance of the Findings'', in the *Dynamics of Automobile Demand*, New York City: General Motors Corporation, pp.3-20.

Hudson, Edward A. and Dale W. Jorgenson [1974], ''U.S. Energy Policy and Economic Growth, 1975-2000'', *The Bell Journal of Economics and Management Science*, Vol. 5, No. 2, Autumn, pp.461-514.

Hulten, Charles R. and Frank C. Wykoff [1981a], ''The Measurement of Economic Depreciation'', in Charles R. Hulten, ed., *Depreciation, Inflation, and the Taxation of Income from Capital*, Washington, D.C.: The Urban Institute Press, pp.81-125.

———— [1981b], ''The Estimation of Economic Depreciation Using Vintage Asset Prices: An Application of the Box-Cox Power Transformation'', *Journal of Econometrics*, Vol. 15, No. 3, April, pp.367-396.

Jorgenson, Dale W. [1974], ''The Economic Theory of Replacement and Depreciation'', Ch. 10 in W. Sellekaerts, ed., *Econometrics and Economic Theory: Essays in Honour of Jan Tinbergen*, London: Macmillan, pp.189-221.

__________ and Lawrence J. Lau [1975], ''The Structure of Consumer Preferences'', *Annals of Social and Economic Measurement*, Vol. 4, No. 1, January, pp.49-101.

Kahn, James A. [1982], ''Gasoline Price Expectations and the Used Automobile Market, 1972-81: Econometric Evidence'', Mimeo, M.I.T. Department of Economics, September.

Lady, George M. and David Nissen [1968], ''Functional Structure in Demand Analysis'', Paper presented at the Winter Econometric Society Meetings, Washington, D.C., December.

Lau, Lawrence J. [1982], ''The Measurement of Raw Material Inputs''. Ch. 6 in V. Kerry Smith and John V. Krutilla, eds., *Explorations in Natural Resource Economics*, Baltimore: Johns Hopkins Press for Resources for the Future, Inc., pp.167-200.

Lucas, Robert E.B. [1975], ''Hedonic Price Functions'', *Economic Inquiry*, Vol. 13, June, pp.157-178.

McFadden, Daniel [1978], ''Cost, Revenue and Profit Functions'', Ch. I.1 in Melvyn Fuss and Daniel McFadden, eds., *Production Economics: A Dual Approach to Theory and Applications*, Amsterdam: North-Holland Publishing Company, Vol. 1, pp.3-109.

Morrison, Catherine J. [1982], ''Three Essays on the Dynamic Analysis of Demand for Factors of Production'', Unpublished Ph.D. dissertation, University of British Columbia, August.

Muellbauer, John [1971a], ''Testing the 'Cagan-Hall' and the 'Hedonic' Hypotheses''. Paper presented at the Summer Econometric Society Meeting, Boulder, Colorado, 27 August, Warwick Economic Research Paper 18, August.

——— [1971b], "The Theory of True Input Price Indices", University of Warwick Economic Research Paper 18, November.

——— [1972], "Characteristics, Substitution Between Goods and Quality", University of London Birkbeck College, Mimeo, November.

——— [1974a], "Testing the Simple Repackaging Hypothesis of Quality Change", University of London, Birkbeck College, Department of Economics, Discussion Paper No. 34, December.

——— [1974b], "Household Production Theory, Quality and the 'Hedonic Technique'", *American Economic Review*, Vol. 64, No. 6, December, pp.977-994.

——— [1975a], "Quality Change in Wheeled U.S. Farm Tractors", Ch. 8 of draft book manuscript, University of Warwick, pp.134-157.

——— [1975b], "The Cost-of-Living and Taste and Quality Change", Journal of Economic Theory, Vol. 10, pp.269-283.

hta, Makota [1975], "Production Technologies of the U.S. Boiler and Turbo Generator Industries and Hedonic Price Indexes for their Products: A Cost Function Approach", *Journal of Political Economy*, Vol. 83, No. 1, February, pp.1-26.

——— and Zvi Griliches [1975], "Automobile Prices Revisited: Extensions of the Hedonic Hypothesis", in Nester Terleckyj, ed., *Household Production and Consumption*, New York: Columbia University Press, pp.325-390.

indyck, Robert S. and Julio J. Rotemberg [1982], "Dynamic Factor Demands, Energy Use, and the Effects of Energy Price Shocks", M.I.T. Energy Laboratory Working Paper No. 82-024WP, April, forthcoming, *American Economic Review*.

ollak, Robert A. [1983], "The Treatment of 'Quality' in the Cost-of-Living Index", forthcoming, *Journal of Public Economics*.

Rosen, Sherwin [1974], "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition", *Journal of Political Economy*, Vol. 82, No. 1, January, pp.34-55.

Roy, R. [1943], *De l'Utilite: Contribution à la Theorie des Choix*, Paris: Hermann et Cie.

Spady, Richard H. [1979], *Econometric Estimation of Cost Functions for the Regulated Trucking Industry*, New York: Garland Press.

__________ and Ann F. Friedlaender [1978], "Hedonic Cost Functions for the Regulated Trucking Industry", *Bell Journal of Economics*, Vol. 9, No. 1, Spring 1978, pp.159-179.

Stone, Richard [1956], *Quantity and Price Indexes in National Accounts*, Paris: Organization for European Economic Cooperation.

Triplett, Jack E. [1971a], "Determining the Effects of Quality Change on the CPI", *Monthly Labor Review*, May, pp.27-32.

__________ [1971b], "Quality Bias in Price Indexes and New Methods of Quality Measurement", Ch. 6 in Zvi Griliches, ed., *Price Indexes and Quality Change*, Cambridge: Harvard University Press, pp.180-214.

__________ [1976], "Consumer Demand and Characteristics of Consumption Goods", in Nester E. Terleckyj, ed., *Household Production and Consumption*, NBER Conference on Research in Income and Wealth, Vol. 40, New York: Columbia University Press, pp.305-24.

__________ [1983], "Concepts of Quality in Input and Output Price Measures: A Resolution of the User-Value Resource-Cost Debate", in Murray F. Foss, ed., *The U.S. National Income and Product Accounts*, Vol. 47 of Studies in Income and Wealth, Chicago: The University of Chicago Press for NBER, pp.269-311.

Wang Chiang, J.S. [1981], "Economies of Scale and Scope in Regulated Industries: A Case Study of the U.S. Trucking Industry", Massachusetts Institute of Technology, Department of Civil Engineering, Unpublished Ph.D. dissertation.

_________ and Ann F. Friedlaender [1982], "Output Aggregation, Network Effects, and the Measurement of Trucking Technology", Mimeo, Massachusetts Institute of Technology, Department of Economics, September.

Waugh, F.V. [1929], *Quality as a Determinant of Vegetable Prices: A Statistical Study of Quality Factors Influencing Vegetable Prices in the Boston Wholesale Market*, New York: Columbia University Press.

COMMENTS

*Jack E. Triplett**
Office of Research and Evaluation
United States Bureau of Labor Statistics

As usual, Ernie Berndt's paper shows a scholarly assimilation of the body of literature he confronts, in this case on the subject of quality change and hedonic methods. However because the hedonic function's relation to the utility function has so often been misunderstood, the literature has often been unclear or misleading concerning what utility theory can say about the form of hedonic functions, and accordingly about the interpretation of "hedonic price indexes." My views on these issues, as presented in the following are compatible with Pollak [1983] and Rosen [1974], as well as, obviously, Triplett [1976 1983a], in opposition to the interpretations in the literature Berndt follows, particularly Muellbauer [1974] and Lucas [1975].

The "Constant-Quality" Cost-of-Living Index

Conventional consumer and index number theory implicitly applies to a world of homogeneous goods. The only analyses of heterogeneous goods yielding measurable implications are those making explicit use of the characteristics of goods. In this, I judge Berndt and I do not disagree: His examples are sprinkled with lists of characteristics, and in his first sentence he specifies that characteristics "affect consumers' satisfactions."

Many authors have taken the approach Berndt adopts in his paper--assuming that some scalar measure of quality exists and that the characteristics explain (in the statistical sense) this scalar measure. I prefer an agnostic position on whether any unique scalar quality measure exists: It is a very special assumption which is not really necessary for the analysis and which tends to obscure relationships that involve the characteristics. Moreover, if the characteristics are the utility-generating quantities in which the consumer is interested, it seems only natural to write them in the utility function, rather than inserting some function of the characteristics in the utility function, as Berndt does.

The general case, then, begins from a utility function defined on the characteristics, i

he manner pioneered by Lancaster [1971]. For simplicity, suppose only one good (the j^{th} one) is non-homogeneous, and that it has m characteristics. We can then write

$$U = U(X_1, X_2, \ldots, Z_{j1}, Z_{j2}, \ldots, Z_{jm}), \quad (1)$$

o that the utility function contains not j goods, but j + m-1 arguments.

To further simplify the problem, suppose both the direct and indirect utility functions re separable on the characteristics of good j. We refer to this as "J-separability." Pollak 1975] and Blackorby and Russell [1978] show that what I have designated j-separability ermits forming a **subindex** of the full cost-of-living index on the characteristics set $[Z_j]$ f, for some "branch utility indicator," W, the minimum cost of characteristics sufficient o attain utility W* under time t conditions is

$$C_t^* = C(W^*, P_j^t[Z_j]) = \min(P_j^t[Z_j] \mid W[Z_j] \geq W^*), \quad (2)$$

ıen the cost-of-living subindex is constructed from a ratio of values from this characteristics cost function" or

$$COL_j = C_t^*/C_o^* \cdot \quad (3)$$

quation (3), the constant-utility subindex on characteristics, is a natural interpretation r what we mean by a "constant quality price index" for good j.

Before proceeding, a few observations about equation (3) are in order. First, we have fined a "constant-quality" price index without any recourse to a "measure" of quality.[1] ıat is a great advantage for empirical work, for the "yardstick" capable of extracting scalar quality indicator from goods has never been discovered. Secondly, the constant-ality price index takes on a form that is a simple translation of normal index number eory into characteristics space. This permits us to look at the theoretical measure with ights gained from analogous analysis on the goods-space case, and dispenses with the ams of special terminology, opaque examples, and "empty boxes" that seem endemic

to alternative approaches. To be sure, the advantages are not purchased without some costs Additional discussion of characteristics-space "input cost" indexes (the cost-of-living in dex belongs to the input cost index class) is in Triplett [1983a].

Utility Functions and Hedonic Functions

It is well known that the hedonic function

$$P_j = P_j\,[Z_j] \tag{4}$$

yields an estimated set of m implicit prices, $\partial P_j/\partial Z_{ji}$, $i = 1,\ldots,m$, for the elements of $[Z_j]$ In much of the literature, the hedonic function of equation (4) has been identified wit the characteristics cost function of equaton (2). If this were correct, then utility theor could be used to specify properties of the hedonic function, along the line of reasonin adopted by Muellbauer [1974], and followed by Berndt. To see the flaw in this reasoning it is worth taking a short digression to review how the cost function in **goods** space is derivec

By duality theory, a correspondence exists between the form of the utility (or produc tion) function and the consumption (production) cost function. However, a key elemen facilitating the correspondence is the parametric nature of prices faced by the individua consuming unit. When prices are fixed regardless of quantities consumed, the budget cor straint takes the form of a hyperplane. In effect, one can think of the consumption co function as the utility function translated by the budget hyperplane. Being able to assum that the boundary of the consumption opportunity set is linear in all relevant dimention greatly simplifies the analysis – in fact, in goods space one seldom thinks about the forn of the function relating expenditure and market prices.

In characteristics space theory, on the other hand, the shape of the budget surface ha from the beginning been perceived as a crucial element (Lancaster, [1971]). There is n reason to assume it is necessarily a hyperplane, though it might be. The shape of the budge surface is an empirical question that depends on the relation between the price of nor homogeneous goods, or of **varieties** of non-homogeneous goods, and the quantities c characteristics embodied in them. This is precisely the information contained in the hedoni

function (equation (4)).

In particular, for the j-separability case described above (and only for that case), the characteristics-space budget surface for the ''branch'' utility function is determined entirely from the hedonic function for good j.[3] To obtain the consumption cost function (equation 3), one combines the hedonic function, as a representation of the budget surface, with the ''branch'' utility function on characteristics, in a manner analogous to the solution of the same problem in goods space. In general, the form of the consumption cost function, and thus of the cost-of-living index, depends on **both** utility function and hedonic function forms. Notice that the form of the utility function does not determine the form of the hedonic function, as the latter is an empirical matter.

The preceding suggests that linearity of the hedonic function in the relevant dimensions would be a useful trait for preserving the simplicity of the analysis. If the hedonic function were to yield a budget surface that had the form of a hyperplane, then the characteristics cost function (equation (2)) – and therefore the characteristics cost-of-living subindex of equation (3) – would have the same straightforward relation to the characteristics utility function as that exhibited by its counterpart in goods space.

On the other hand, if the hedonic function yields a curved budget surface, then computing the cost-of-living index would require knowledge of the form of the utility function, and of the hedonic function, plus a bit of manipulation. With a curved budget surface, the unique correspondence between the form of the utility function and the form of the consumption cost function (and therefore of the index) that is familiar in usual goods-space analysis disappears.

Curvature creates a further problem. Consumers who locate at different portions of the surface face different prices for characteristics. This property of the characteristics world is well documented (it is discussed in different contexts by Lancaster [1971], Rosen [1974] and Pollak [1983]). The need to model ''locational'' parameters greatly complicates any demand estimation in characteristics space (Rosen [1974]; Brown [1983]; Triplett [1983b]). Though the problem is ignored in the work of Lau [1982], on which so much of Berndt

depends, any practical attempt to implement Lau's "conversion ratios" must face up t the locational problem.

To summarize the linearity matter, only if the hedonic function exhibits linearity in th relevant dimensions will the characteristics space cost-of-living index number assume form normally thought to be consistent with those of "flexible" utility functions. Berndt's - and Muellbauer's -- statements about the compatibility of the characteristics space inde number and the utility function are well taken, but are misdirected: The implications the draw properly concern the form of the characteristics cost function (but even then are cor rect only under budget surface linearity). The correspondence between utility function an index number forms in characteristics space is closer the more nearly the consumption op portunity locus approximates a hyperplane, and that requires a hedonic function linea in the relevant dimensions. But of course theoretical convenience cannot assure the ap propriate linearity exists, empirically, and the theory cannot determine the shape of th budget surface, which is an empirical matter.

I have several times attached the term "relevant dimensions" to a reference to hedoni function linearity. What matters for present purposes, and for demand analysis, is nc whether the hedonic function **itself** is linear, but the shape of the budget surface it yield The semi-log hedonic function is an example of a non-linear funcitonal form that is neve theless linear in the "relevant dimensions" and accordingly yields a hyperplane as a budg surface in characteristics space (for the documentation, see Triplett [1976]). The sectio of Berndt's paper in which he (following Muellbauer and Lucas) concludes that th semi-log form is theoretically inferior to other forms of hedonic functions stands th matter exactly on its head: As a representation of the budget surface, the semi-log hedoni function has the nice theoretical property that it yields a hyperplane as the form of th surface.

Empirical tests on hedonic function forms have usually found the semi-log to provic a good fit to the data.[4] This is one case in which it is good thing for the theory that th empirical world has confirmed (tentatively) a property that some theorists said it shouldn have!

I do not wish to minimize the complexity of characteristics space analysis. In the preceding, ›nly the quantity of characteristics was assumed to matter, and not the quantity of goods nto which the characteristics were packaged. This might be true of soap or cornflakes where the numer of ounces in the package is one characteristic), but is unlikely for more omplex goods (as Trajtenberg [1979] so colorfully put it, ''Two fiddles don't make a itradivarius''). Yet, my intuition says that for the restrictive j-separability case the assumpion that only quantities of characteristics matter is not additionally restrictive.

-separability and the Existence of a Scalar Quality Measure

I have noted already that the characteristics-space analysis of quality makes a scalar ıeasure unnecessary. Under the characteristics approach, ''quality'' can be interpreted s merely economic shorthand for the quantities in a vector of characteristics.

The characteristics-space approach can also show when a unique scalar measure of quality ın be formed, and be used to illuminate the properties and usefulness of a scalar measure. erndt's scalar measure "b", inserted into the utility function of his equation (1), can be ıought of as an aggregation over the characteristics included in his equation (2) – or in ıe notation of this comment, an aggregation of the m characteristics of good j. Tradional aggregation theory makes clear that such an aggregation yields a unique value only ›r what we have here termed j-separability. Indeed, j-separability is an underlying imlicit assumption in nearly all the literature that treats ''quality'' as some scalar arameter,[5] though acknowledged in almost none of the literature which takes this aproach. It occurs in the analysis of labor ''quality'' as well, for example in the work of hinloy [1980], as well as Gallop and Jorgenson [1983] and in other work on human capital.

Because the complexity of characteristics-space analysis seems a deterrent to its use, there great appeal to simpler cases. The difficulty is that what seem to be simpler approaches, ıch as Berndt's scalar ''b'' and Lau's ''conversion ratios,'' rest on j-separability. If the ıaracteristics world is more complex (and it undoubtedly is), approaches using scalar easures may yield invalid results, and at best are approximations that obscure the real ıture of the problem.[6]

The assumption of j-separability is not a very realistic one. Automobile characteristics are probably not separable from the price of gasoline, nor those of refrigerators from ice cream. When j-separability does not hold, scalar measure approaches to the analysis of quality are problematical.

Moreover, the connection between utility function and index number forms becomes even more remote than in the j-separability case, because the budget surface for a utility branch is not identical with the surface defined by the hedonic function. Instead, the con-straint must be computed from prices (implicit and otherwise) of all goods included in the branch. When more than one non-homogeneous good is included in a branch--suppose for example, that gasoline has characteristics such as octane rating and so forth, which must be incorporated into a branch containing automobile characteristics as well--then the budget surface is some combination of the separate hedonic functions. The consumption cost function (equation 2) is, as before, composed from the utility function and the budge surface. I presume that some program could be worked out for this computation, but to my knowledge no one has much of an idea what such a construction would look like, and the information for solving even fairly simple cases has not been assembled.[7]

On Interpreting ''Hedonic Price Indexes''

One remaining point requires comment, among others that could be explored, space per-mitting: What are those ''hedonic price indexes'' that researchers have computed by throw-ing time dummy variables into multi-period hedonic functions?[8] As evident from th above discussion, they are not characteristics-space cost-of-living indexes. Moreover, th imposition of flexible utility function forms on them would do nothing to move them towar the true indexes.

Under the j-separability assumption, the hedonic function is interpretable as a direc estimation of budget surface.[9] Estimating an index number directly from the hedoni regression thus amounts to computing an index number from the budget surface, withou use of information from the utility function. Analogous computations are commonly car-ried out in goods space--ordinary fixed-weight indexes, such as Laspeyres or Paasche for-mulas, likewise use only information from the budget hyperplane. They are interprete

in the index number literature as approximations to the true cost-of-living index since the latter combines information from both the budget hyperplane and the utility function. Ac- :ordingly, it seems reasonable to give a similar interpretation to ''hedonic price indexes'' :omputed directly from regressions--they are approximations, based on information about .he characteristics-space budget surface, to the true characteristics space cost-of-living in- lex. Owing to the character of the approximation, the most one should require of the ıedonic function is that it provide a good fit as an empirical realization of the budget surface.

The role of theory in the characteristics-space problem should be similar to its employ- nent in the standard goods-space analysis. Theoretical work on characteristics space cost- ıf-living indexes has barely begun (see Pollak, [1983] and Triplett [1983a]), and many of he properties of such indexes are not worked out. It may turn out that, just as Laspeyres rice indexes seem fairly close approximations to the goods-space cost-of-living index (see 3raithwait, [1980]); hedonic price indexes computed directly from regressions are fairly ood approximations to characteristics-space cost-of-living indexes. But one cannot ''im- rove'' the approximation without knowing the true index and its properties. Theoretical 'ork needs to be directed toward uncovering the properties of the true index in haracteristics space, and not misdirected, as so much of it has been in the past, to trying) find functional forms for hedonic functions that ''satisfy'' axioms of utility theory. 'hat is chasing a will-o'-the-wisp.

ummary

There are four separate characteristics-space functions that have often become confus- l in the literature: the hedonic function, the budget surface, the consumption cost func- on, and the utility function. Only under what I have termed j-separability do the first vo of these functions correspond. Only when the hedonic function yields a hyperplane ır the budget surface is there the kind of straightforward correspondence between the :ility function and the consumption cost function that parallels the one that exists in goods ›ace. But not even with j-separability, which is clearly the simplest case, is the form of .e hedonic function derivable from the utility function. The shape of the budget surface, ıd accordingly of the hedonic function, is purely an empirical matter that is appropriate- resolved by seeking the best fit to the data. In particular, nothing in consumption theory

rules out the well-worked semi-log functional form for hedonic functions — quite to the cor
trary, for the semi-log hedonic function, if it conforms to empirical reality, turns out t
be a convenient functional form, theoretically.

An additional analytical point is that j-separability, restrictive and unrealistic as it is is nevertheless an implicit assumption that lies behind scalar measure approaches to th analysis of quality change. This suggests that scalar measure analyses are not very general and that the assumption that one can, e.g., write the utility function in terms of the quar tities and quality ''levels'' of goods is a very special case. Accordingly, the scalar assump tion is not a good starting point for theoretical reasoning about heterogeneous goods an quality change.

Footnotes

Associate Commissioner, U.S. Bureau of Labor Statistics. Views are those of the author, and do not represent official positions of the Bureau of Labor Statistics or the U.S. Department of Labor.

Except for the branch utility indicator, W. This, however, is unique to the j-separability case. Relaxation of the separability assumption is discussed below.

Writing the hedonic function this way makes it clear that the variables in hedonic functions should in principle be chosen so that they are arguments of the utility function (when we are talking abut the consumption case). This rule has not been followed very closely in most of the existing hedonic studies, in which the independent variables are best interpreted as proxies for the true utility-generating characteristics.

Suppose U is not separable on the characteristics of j. Then the budget surface will contain prices of other goods, and those may not combine with the characteristics prices in any simple fashion. It remains true that the hedonic function on good j provides information for computing the surface, but it may be a complex computation. See the further discussion, below, note 6.

Unfortunately, our confidence in this finding is diminished by the fact that the empirical testing was either bereft of any theory at all or influenced by the wrong one, with the results that some otherwise plausible functional forms have not been considered at all.

A notable exception is the early work of Houthakker [1952], which is often appealed to by writers on hedonic functions. Houthakker explicitly considered the case where a good had only one "quality indicator" (which I interpret to mean a single characteristic in modern terminology), so his framework does not require the j-separability assumption. At the end of his article, Houthakker cautions that the case of multiple characteristics requires a different and more complex theory, a warning not heeded by the various later writers who have tried to link multiple-characteristics analysis to the Houthakker article.

Note in particular that in Lau's conversion ratio, as cited in Berndt's equations (4) and (5), the existence of the unique scalar quality measure, b_n, rests on j-separability, but the conversion ratio itself depends on quantities of all the consumption goods, which dependence is trivial if separability holds. This dependence is acknowledged in later equations, and the implication of something very like j-separability is discussed, but nowhere is it acknowledged that the general case in which the conversion ratio depends on all goods means that the quality measures included in the conversion ratios are not unique scalars.

Berndt cites several unpublished studies that have added the price of gasoline to a hedonic function on automobile characteristics. I suppose one might rationalize these as attempts to estimate the budget surface directly, or--more problematical--as approximations to the consumption cost function.

The most common form of this calculation (see Griliches, [1971]) is $\ln P = Xb + ct$ where t is an indicator variable for the second of two, or for the i th of n, time periods, and $1 - \exp c$ is interpreted as the percentage change in the index.

The slopes of the budget surface are given by $\partial Z_r / \partial Z_s$, P_j = constant.

Reference

Blackorby, Charles and R. Robert Russell, [1978], Indexes and subindexes of the cost-of-living and standard of living. *International Economic Review* 19 (February), pp.229-240

Braithwait, Steven D. [1980], The substitution bias of the Laspeyres price index: An Analysis using estimated cost-of-living indexes. *American Economic Review* 70, pp.64-77.

Brown, James N. [1983], Structural Estimation in Implicit Markets. In Jack E. Triplett, ed. *The Measurement of Labor Cost*. Conference on Research in Income and Wealth, Studies in Income and Wealth, Volume 48. Chicago: University of Chicago Press for the National Bureau of Economic Research. pp.123-151.

Chinloy, Peter, [1980], Sources of quality change in labor input. *American Economic Review* 70 (March), pp.108-119.

Gallop, Frank M. and Dale W. Jorgenson, [1983], Sectoral measures of labor cost for the United States, 1948-1978. In Jack E. Triplett, ed., *The Measurement of Labor Cost*. Conference on Research in Income and Wealth, Studies in Income and Wealth, Volume 48. Chicago: University of Chicago Press for the National Bureau of Economic Research pp.185-235.

Houttakker, Hendrick S. [1952], Compensated changes in quantities and qualities consumed *Review of Economic Studies* 19: pp.155-164.

Lancaster, Kelvin, [1971], *Consumer Demand: A New Approach*, New York: Columbia University Press.

Lau, Lawrence J. [1982], The measurement of raw material inputs. In V. Kerry Smith and John V. Krutilla, eds., *Explorations in National Resource Economics*. Baltimore: Johns Hopkins University Press for Resources for the Future.

Lucas, R.E.B. [1975], Hedonic price functions. *Economic Inquiry* 13: pp.157-178.

/luellbauer, John, [1974], Household production theory, quality and the ''hedonic technique.'' *American Economic Review* 64: pp.977-994.

'ollak, Robert A. [1975a], Subindexes of the cost of living. *International Economic Review* 16: pp.135-150.

——— [1983], The treatment of ''quality'' in the cost-of-living index. Forthcoming in the *Journal of Public Economics*.

.osen, Sherwin, [1974], Hedonic price and implicit markets: Product differentiation in pure competition. *Journal of Political Economy* 92: pp.34-55.

'rajtenberg, Manuel [1979], Quality is all very well but two fiddles don't make a Stradivarius. Falk Institute discussion paper.

riplett, Jack E. [1976], Consumer demand and characteristics of consumption goods. In Nestor E. Terleckyj, *Household production and consumption,* Conference on Research in Income and Wealth, Studies in Income and Wealth, Volume 40. New York: National Bureau of Economic Research.

——— [1983a], Concepts of quality in input and output price measures: A resolution of the user value-resource cost debate. In Foss, Murray, ed., *The U.S. national income and product accounts: selected topics*. Conference on Research in Income and Wealth: Studies in Income and Wealth, Volume 47. University of Chicago Press for the National Bureau of Economic Research.

——— [1983b], An essay on labor cost. In Jack E. Triplett, ed. *The measurement of labor cost*. Conference on Research in Income and wealth, Studies in income and wealth, volume 48. Chicago: University of Chicago Press for the National Bureau of Economic Research.

MEASURING THE CURRENT RATE OF INFLATION

Pour vous fournir une version dans la langue officielle de votre choix, le texte anglais est suivi du texte français (p.900) dans cette publication.

D. Rhoades and N. Elhawary-Rivet
Current Economic Analysis Division
Statistics Canada

SUMMARY

The goals of the paper are:

a) To arm the reader with sufficient knowledge to use both the annual and monthly rates to maximum advantage in interpreting inflationary developments.

b) To explore new ways of monitoring the cyclical evolution of inflation that attempt to utilize information more efficiently, and thereby give a more accurate picture of the economy.

It is shown that the annual percentage change of a time series is a smoothed version f the monthly percentage change. It is also shown that the annual percentage change lags ie monthly as it shifts the data 5.5 months forwards relative to the monthly percentage ange. Both intuitive and technical approaches are given in Sections II and III respectively.

The search for alternative measures of the rate of inflation revolves around reducing e delay inherent in the year-over-year rate while maintaining the smoothing it achieves. he properties of two alternatives are presented. In the first alternative, proposed by eoffrey Moore, the price level in the current month is compared to the average price level er the immediately preceding 12 months. The second alternative provided by Rhoades, nploys minimum phase shift filtering to smooth the monthly percentage change series. iese methods produced smoothed series with smaller time lag than the annual percentge change but retained some seasonal cycles which existed in the original series.

To illustrate the discussion, two practical examples using the ISPI and CPI were given. The three different methods were applied to the time series and the results were compared in terms of smoothness and time lag.

I Introduction

In a recent editorial on inflation the Financial Times of Canada (FT 2/8/82) called upon the government to "devise ways to tell Canadians exactly where we stand. We now have an inflation rate that seems to go up and down at the same time. If Canadians are expected to join the crusade, they need to know exactly how strong the enemy is; instead of being flim-flammed with statistical tricks."

Much of the problem of the Financial Times, and presumably of other users, was generated by the use of both monthly and annual rates of change in presenting the overall rate of inflation. Confusion is created by the fact that the monthly rate may decline at the same time as the annual rate increases. Knowledge of the relationship between the monthly and annual rates of change enables reconciliation of these seemingly contradictory movements. Thus, one goal of this paper is to arm the reader with sufficient knowledge to use both the annual and monthly rates to maximum advantage in interpreting inflationary developments.

A second goal of the paper is to explore new ways of monitoring the cyclical evolution of inflation that attempt to utilize information more efficiently, and thereby give a more accurate picture of where we stand. These alternate measures impose, however, an even greater expository responsibility on the Statistical Agency, and the meeting of that responsibility is a final goal of this paper.

Sections II and III on the properties of the monthly (month over preceding month) and annual (month over same month a year ago) percentage changes make three major points:

(i) The annual percentage change is a smoothed version of the monthly percentage change.

ii) All smoothing schemes shift the smoothed signal forward in time, causing a delay in recognizing cyclical changes.

iii) The annual percentage change lags the monthly as it shifts the data 5.5 months forward relative to the monthly percentage change.

Section II presents the arguments in intuitive terms, while Section III makes the same points in technical terms. Readers not so inclined may skip Section III without losing the essence of the argument.

The search for alternative measures of the rate of inflation revolves around reducing the delay inherent in the year-over-year rate while maintaining the smoothing it achieves.

An alternative has been proposed by Geoffrey Moore [2] in which the price level in the current month is compared to the average price level over the immediately preceding 12 months. Another alternative is provided by Rhoades [4] which employs minimum phase shift filtering to smooth the monthly percentage change series. The properties of both these alternatives are reviewed in Section IV.

I The Intuitive Approach

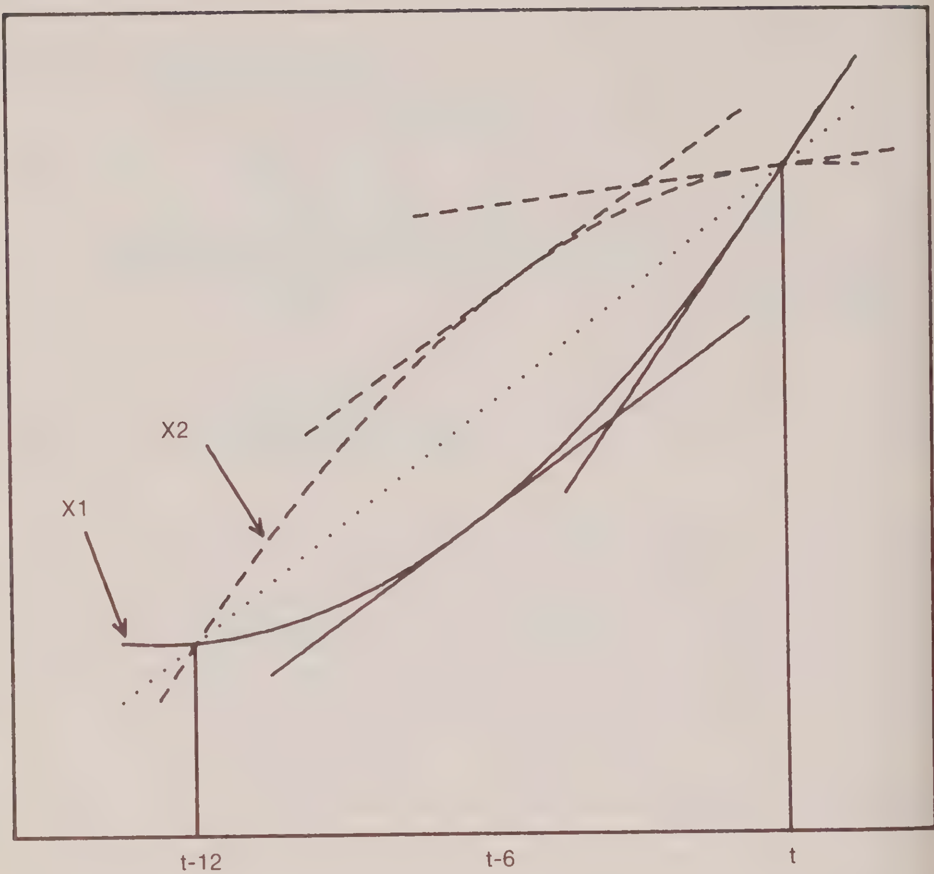
In this section we discuss in intuitive terms the properties of the annual and monthly percentage changes. The first point we wish to establish is that the annual percentage change is a smoothed version of the monthly percentage change. This is clear from the following:

$$\frac{x_t}{x_{t-12}} = \frac{x_t}{x_{t-1}} \cdot \frac{x_{t-1}}{x_{t-2}} \cdots\cdots \frac{x_{t-11}}{x_{t-12}}$$

$$\log \frac{x_t}{x_{t-12}} = \sum_{k=1}^{12} \log \frac{x_{t-k+1}}{x_{t-k}}$$

which shows that the annual percentage change is approximated by a moving average of the monthly. Consider the following example to illustrate some further properties of the two measures.

Figure 1



Let x_1 and x_2 be two time series in Figure 1 below.

The annual percentage change is approximated by the slope of the line connecting x_1(t) and x_1(t-12), or x_2(t) and x_2(t-12). In this case, since x_1(t) and x_2(t), and x_1(t-12) and x_2(t-12) coincide, the annual percentage change is the same for both series. The monthly percentage change at time t, however, is approximated by the slopes from t-1 to t, and is positive for x_1 but negative for x_2.

Two important points can be shown from Figure 1:

(i) Because the annual percentage change ignores the path taken in moving from t-12 to t it is unable to discriminate the direction of motion at time t.

(ii) The slope of both curves at time t-6 approximates the annual percentage change. This means the annual percentage change represents the monthly change six months ago (i.e. it lags the original series six months).

These differences are caused by the fact that the annual percentage change depends on x only at times t and t-12 and hence:

(a) It ignores the information in the intervening months of t and t-12.

(b) It projects any peculiarities and irregularities that existed a year ago into the present.

III The Theoretical Approach

In this section we discuss more rigorously the theoretical differences between the annual and monthly percentage changes. To compare the two we will consider the gain G(f) and phase P(f) functions [1] associated with each. The gain and phase functions for the monthly percentage change (see Appendix 1) are

$$G_1(f) = \left(\frac{100}{x_{t-1}}\right)^2 4 \sin^2 \pi f \qquad (0<f\leq .5)$$

$$P_1(f) = -.5 + \frac{1}{4f} \quad \text{in months} \quad (0<f\leq .5)$$

where f = frequency.

The gain and phase functions for the annual percentage change filter are

$$G_{12}(f) = \left(\frac{100}{x_{t-12}}\right)^2 4 \sin^2 12\pi f \qquad (0<f\leq.5)$$

$$P_{12}(f) = -6 + \frac{1}{4f} \qquad \text{in months} \quad (0<f\leq.5)$$

It is obvious that the two percentage changes have different gain functions, but they both have similar form. Each gain is a function of a sine wave which has different periodicity in each case (the period is 2 for the monthly and 1/6 for the annual).

The phase function is of more interest here. The annual percentage change shifts the frequency component f forward in time by about six months (exactly 6-(1/4f) months), whereas the monthly percentage change shifts frequency f forward by .5-(1/4f) of a month.

The difference between these two phase functions gives the displacement in time of annual percentage changes relative to the monthly. This calculation indicates that the annual percentage change series will lag the monthly percentage changes by 5.5 months.

IV An Alternative to the Annual Percentage Change

The previous section showed that the annual percentage change series is a smoothed version of the monthly percentage change series and that it has five and one-half months of phase shift relative to the monthly. This means that to make an accurate assessment for

the current month one has to wait for about six months. It would be better if another method of smoothing the monthly percentage change series could be found with a smaller phase shift. In this section we discuss two alternatives.

The first of these methods is the one used by Rhoades [4] which utilizes spectral analysis techniques to design filters that minimize the phase shift for a given degree of smoothing.

A suitable filter for smoothing the monthly percentage change series was found to be an autoregressive moving average (ARMA) filter with two AR terms and one MA term.[1] The AR coefficients are $b_1 = 1.451$ and $b_2 = -0.5857$, and the MA coefficient is $a_0 = 0.134$.

The second method which was developed by Moore [2] is defined as follows

$$y_t = \frac{x_t - \frac{1}{12}\sum_{k=1}^{12} x_{t-k}}{\frac{1}{12}\sum_{k=1}^{12} x_{t-k}}$$

where y_t is the smoothed series and x_t is the original level series.

The smoothed series y_t in this case can also be expressed as a moving average of the monthly percentage changes, and hence one can use spectral analysis to draw conclusions about the effect of this filter on the monthly percentage change data.

The gain and phase functions (relative to the monthly percentage change) for the ARMA filter, for the annual percentage change filter, and for Moore's filter are plotted in Figures 2 and 3.

At any given frequency, the gain function tells us by how much the amplitude of the filtered series is reduced ($G(f) < 1$) or amplified ($G(f) > 1$), and the phase function tells us by how much the filtered series is shifted forwards or backwards. (The amplitude of a given frequency component in the filtered series is given by the amplitude of that component in the original data times the value of the gain function at that frequency.)

Figure 2
Gain Function

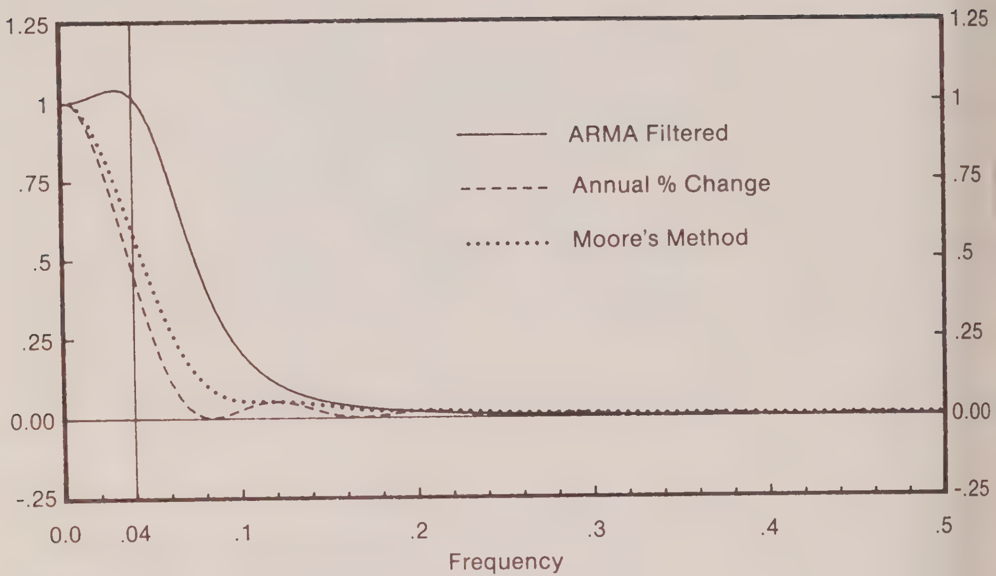
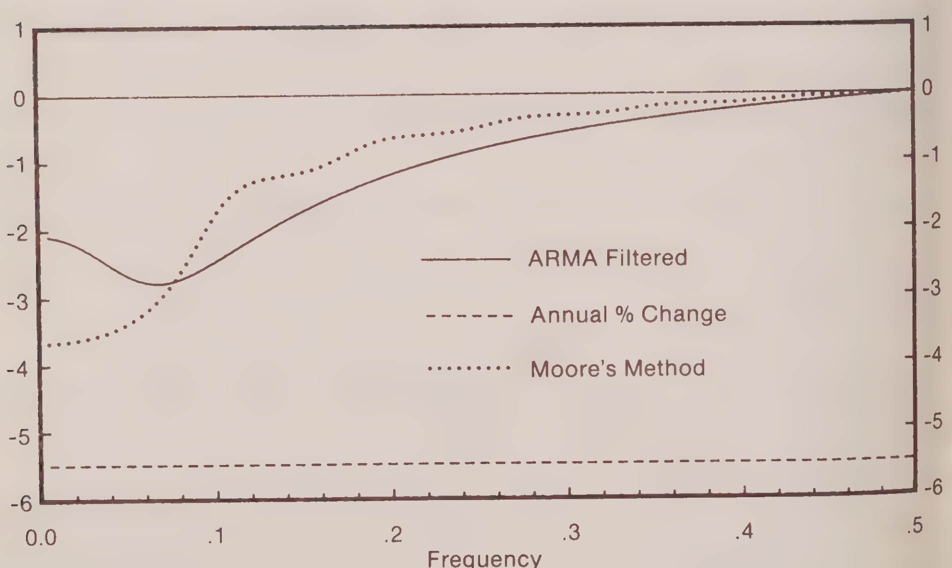


Figure 3
Phase Function



If we take f = 0.04 (periodicity = 25 months) as the dividing line between trend-cycle and seasonal-irregular components we can see that the gain function for the ARMA filter preserves virtually all of the trend-cycle component. On the other hand, the gain functions for the annual percentage filter and Moore's filter do not preserve all the cycles. For example the annual percentage change will retain only 68 per cent of the amplitude at f = 0.03 (periodicity = 33 months), showing that the annual percentage change reduces the amplitude substantially for some cycles. On the other hand the gain function for the ARMA filter retains some of the seasonal and irregular, since over the frequency interval [0.04,0.2] this gain function is not near zero, which means that the amplitude for the seasonal and irregular components will not be reduced to zero at these frequencies.

The phase functions for the three above filters are shown in Figure 3. The average phase shift for the ARMA filter for the frequency interval [0,0.04] is about two months rather han the 5.5 for the annual percentage change filter and 3.5 months for Moore's filter. Therefore, despite some penalties in terms of retaining some of the seasonal and irregular omponents, the ARMA filter has a smaller phase shift while at the same time leaving he trend-cycle component intact. Moore's method has an average phase shift of about .5 months and achieves somewhat less smoothing than the annual percentage change.

The ISPI and CPI annual percentage change series and the filtered monthly percentage eries (using the ARMA filter and Moore's filter mentioned above) are plotted in Figures and 5, illustrating further the earlier comparison of the three gain functions. It is clear hat there is a difference in phase shift as the annual percentage series is shifted more to e right than is the filtered monthly percentage series or Moore's series. The annual percenage change series is somewhat smoother than the others because it removes the seasonal s well as the irregular components, whereas the ARMA and Moore's filter retain some f the seasonal and irregular components. However, much of the additional variance in e other filters appears to be due to seasonal cycles, and therefore these filters are limited displaying movements in seasonally adjusted data. The movement of Moore's data and e ARMA filtered series is very similar, even at the seasonal band of frequencies.

Figure 4
Variation annuelle en pourcentage de l'IPVI
(Janvier 1971 à septembre 1982)

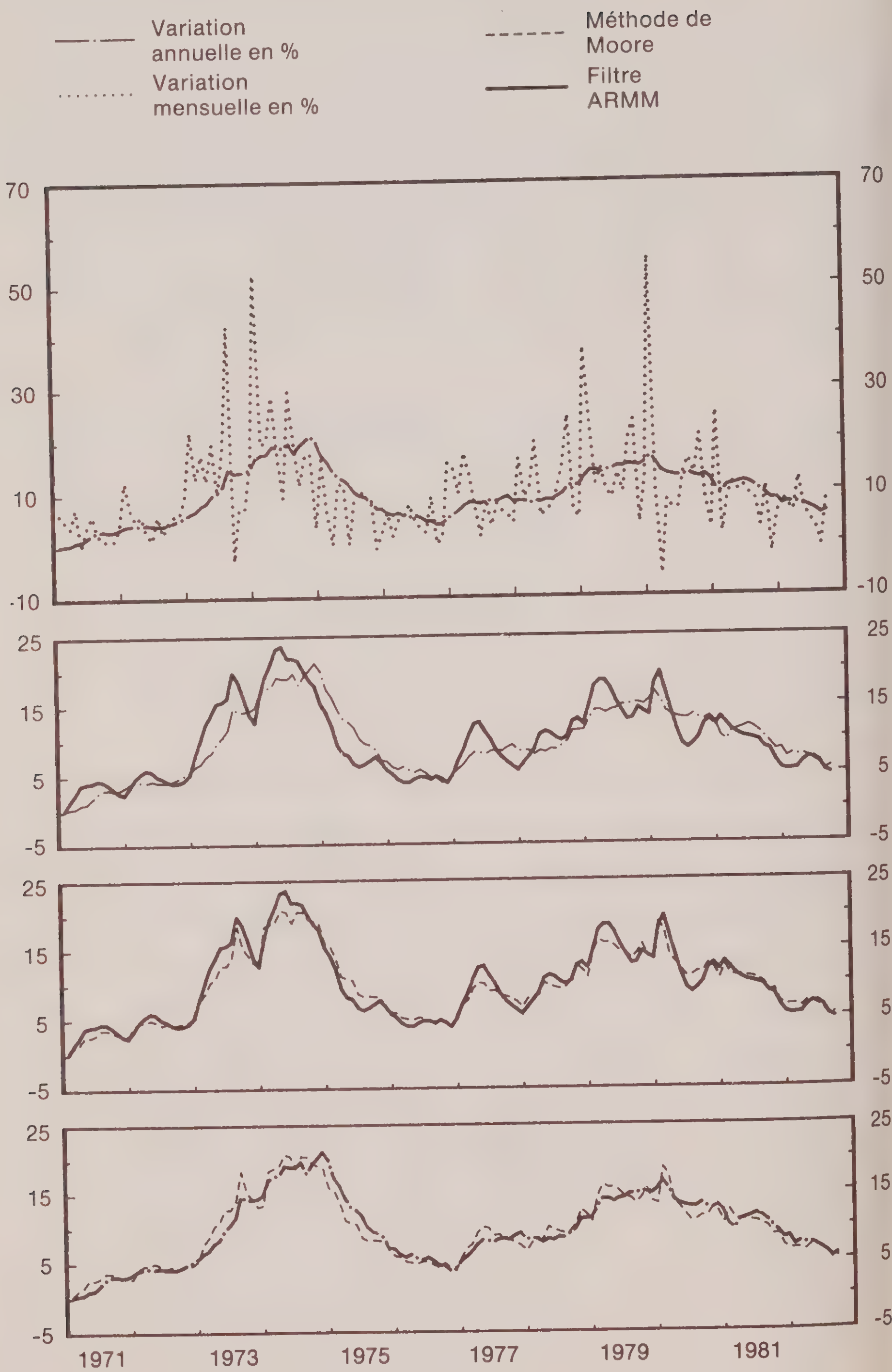
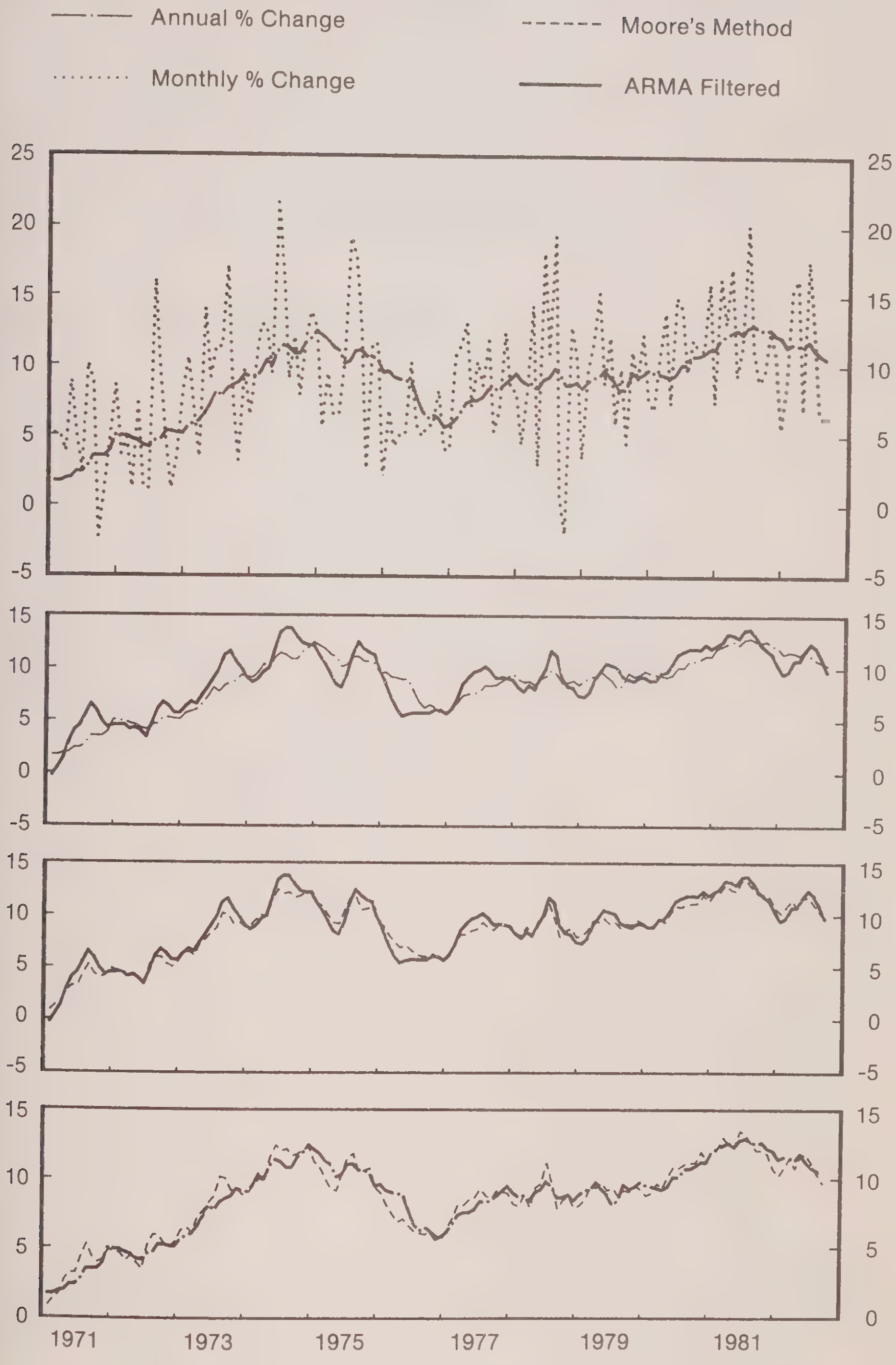


Figure 5

CPI Percentage Change at Annual Rates
(January 1971 to September 1982)



V Conclusion

We have shown that the annual percentage change is a smoothed version of the monthly percentage change series, and that it has 5.5 months phase shift with respect to monthly percentage changes. We have also demonstrated that one could smooth the monthly percentage change series using Moore's method or an ARMA filter, instead of the annual percentage change series. These smoothed series have less phase shift than the annual percentage change but retain seasonal cycles which exist in the original series.

Knowledge of these properties should enable users to reconcile and interpret seemingly contradictory monthly and annual movements in the CPI and other data. For example, in April 1981 the monthly percentage change in the CPI dropped to 0.7 following three months in each of which it was over 1.0. This decline in the CPI had been anticipated by many analysts as inflation in commodity and industrial prices had already moderated in response to the recession in the first half of 1980. In retrospect April also appears to have been the beginning of an easing in inflation rates that has persisted to the present. Unfortunately the annual percentage change rose in April to 12.6 from 12.4 the month before, and this confused many people's perception of events at the time. Much was made in the press of the rise to 12.6 per cent as this shaded, by 0.1 per cent, the previous post-war record. Thus, at a moment when inflation was actually beginning to decline many people thought it was accelerating to new highs because they were using a measure that was almost six months out of date. In order to reduce the possibility of such interpretations, it is recommended that as long as the monthly and annual percentage changes are used, press releases contain a note explaining their relative properties regarding smoothness and timeliness. The question of an alternative to the annual percentage change should be investigated further and a preferred alternative should be selected.

Appendix 1

Transfer Function for the Annual and Monthly Percentage Change Filters

The monthly percentage change

$$y_t = \frac{x_t - x_{t-1}}{x_{t-1}} 100$$

an be viewed as a time varying moving average

$$y_t = a_0 x_t + a_1 x_{t-1} \quad \text{where} \quad a_0 = \frac{100}{x_{t-1}} = -a_1$$

he transfer function [1] of this moving average filter is

$$A_1(f) = a_0 + a_1 e^{-2\pi i f} = \frac{100}{x_{t-1}}(1 - e^{-2\pi i f})$$

$$= \frac{100}{x_{t-1}} e^{-\pi i f} (e^{\pi i f} - e^{-\pi i f})$$

$$= \frac{100}{x_{t-1}} e^{-\pi i f} (2i \sin \pi f)$$

$$= \frac{100}{x_{t-1}} e^{-\pi i f} \cdot e^{\frac{\pi}{2} i} (2 \sin \pi f) \quad \text{where } i = e^{\frac{\pi}{2} i}$$

$$= e^{(\frac{\pi}{2} - \pi f) i} \frac{200}{x_{t-1}} \sin \pi f$$

he gain function $G_1(f)$ is the squared modulus $|A_1(f)|^2$ of the transfer function, and the hase function $P_1(f)$ in radians is the complex angle of $A_1(f)$. Thus, for the monthly percen-ge change

$$G_1(f) = 4\left(\frac{100}{x_{t-1}}\right)^2 \sin^2 \pi f$$

$$P_1(f) = \frac{\pi}{2} - \pi f \qquad \text{in radians}$$

$$\frac{P_1(f)}{2\pi f} = \frac{1}{4f} - .5 \qquad \text{gives the phase displacement in periods.}$$

It is interesting to note that only the gain function varies with time.

An analogous argument yields the gain and phase functions for the annual percentage change.

$$y_t = a_0 x_t + a_{12}\, x_{t-12} \quad \text{where } a_0 = \frac{100}{x_{t-12}} = -a_{12}$$

$$A_{12}(f) = a_0 + a_{12}\, e^{-24\pi i f}$$

$$= e^{(\frac{\pi}{2} - 12\,\pi f)i}\left(\frac{200}{x_{t-12}} \sin 12\pi f\right)$$

$$G_{12}(f) = 4\left(\frac{100}{x_{t-12}}\right)^2 \sin^2 12\pi f$$

$$P_{12}(f) = \frac{\pi}{2} - 12\pi f \qquad \text{in radians}$$

$$\frac{P_{12}(f)}{2\pi f} = \frac{1}{4f} - 6 \qquad \text{in periods.}$$

As pointed out in Section II, the annual percentage change can be viewed as a smoot[h] version of the monthly percentage changes. $A_1(f)$ and $A_{12}(f)$ give us information ab[out] the behaviour of the monthly and annual percentage changes relative to the original d[ata] x. It is also interesting to inquire about the behaviour of the annual relative to the mont[hly]

ercentage change. It can be shown [1, p.45] that the transfer function relating the mon-
hly and annual percentage changes is given by $A(f) = A_{12}(f)/A_1(f)$.

$$\text{Thus } G(f) = \frac{G_{12}(f)}{G_1(f)} = \left(\frac{x_{t-1}}{x_{t-12}}\right)^2 \left(\frac{\sin 12\pi f}{\sin \pi f}\right)^2$$

$$\text{and } P(f) = P_{12}(f) - P_1(f) = -11\pi f \text{ in radians}$$

$$\text{or } \frac{P(f)}{2\pi f} = -5.5 \text{ is the phase shift in months.}$$

his shows that the annual percentage change is a smoothed version of the monthly percen-
ge change that is phase shifted by a constant 5.5 months.

Footnote

1 The filter is of the form $y_t = b_1 y_{t-1} + b_2 y_{t-2} + a_0 x_t$ where x_t is the original data and y_t is the smoothed data.

Reference

[1] Jenkins, G.M. et D.G. watts, *Spectral Analysis and Its Applications,* Holden-Day, 1969.

[2] *Inflation Watch*, Mai-juin 1982, American Enterprise Institute for Public Policy Research.

[3] Nerlove, M., *Spectral Analysis of Seasonal Adjustment Procedures,* Econometrica, Vol. 32, juillet 1964.

4] Rhoades, D., *La conversion de l'actualité en fiabilité des séries chronologiques économiques,* Revue canadienne de statistique, février 1980.

COMMENTS

Stephen B. Browne
Chief Economist
Sun Life Assurance Co. of Canada

Conference participants have stressed the need to ask the right question. The authors address an important one: "What is the current rate of inflation?" The importance o this question derives from its impact on the formulation of inflationary expectations whicl play a pervasive and critical role in economic behaviour and macroeconomic policy formulation.

The paper deals more specifically with four related questions, the answers to which repre sent approximate responses of varying quality to the principal question. These questions are

(i) How much higher is the price index this month than it was in the corresponding mont one year earlier?

(ii) How much higher is the price index this month than last month?

(iii) How much higher are prices now than they were on average during the past year

(iv) What is the current rate of change as determined using spectral analysis techniqu to design a filter that minimizes the phase shift for a given degree of smoothing

The answer to (i) is provided by the 12-month percentage change which is common and widely used as the measure of the current rate of inflation. As correctly documente by the paper, its desirable smoothing properties come at the cost of timely recognition movements in current inflation rate – i.e. it has a phase shift of 5.5 months which mea "... that to make an accurate assessment for the current month one has to wait for abo six months." Question (ii), or the month-to-month percentage increase, is expected to yi volatile and misleading signals owing to the presence of a high degree of "noise" contai ed in the monthly data. The answer to question (iii) is provided by what the authors re to as Moore's filter, which was recently proposed[1,2] as a smoothing technique appropri

to monthly economic time series in general, not solely prices. The answer to question (iv) s provided by the authors in the form of a minimum phase shift autoregressive moving ıverage (ARMA) process.

As demonstrated in the paper, the answers to questions (i), (iii) and (iv) can "be ex-)ressed as a moving average of the monthly percentage change, and hence one can use pectral analysis to draw conclusions about the effect of this filter on the monthly percen-age change data." The principal conclusion drawn from the investigation using spectral nalysis techniques is that "... The ARMA filter is to be preferred on grounds of a smaller hase shift" (2 months versus 3.5 months for Moore's filter and 5.5 months for the an-ual percentage change filter). However, given very similar movements of the ARMA and Aoore filters, the former, in the view of the authors, has only a "slight edge" over the latter.

There are at least five major issues raised by the paper, four of which need to be fully ‹plored before any concrete proposal for a preferred measure of the current rate of infla-on could be advanced. The fifth points out that the policy maker and the expectational ·ocess will need information in addition to whatever measure of the current inflation rate decided upon. These five issues, which will be addressed in turn, are: the choice of an dex; seasonal adjustment; the ability of users to readily verify the preferred algorithm ɔm published data; the full examination of the larger set of sub-annual filters which could principle meet the selection criteria; and, the remaining need to address the question a so-called "core" or "underlying" rate of inflation.

The authors apply their analysis to two price series – the Industrial Selling Price Index SPI) and the Consumer Price Index (CPI). No guidance is provided as to which of the ries, if either, should be employed in search of the answer to the general question: "What the current rate of inflation?" It may well be, for example, that the implicit price deflator Gross National Expenditure or personal consumption may provide a more appropriate se series from which to work. Future work will need to sort out this issue.

It is general practice to base sub-annual computations on data from which the seasonality s been removed (where "seasonality" is meant to include those sub-annual processes ich are regular in occurrence although they may not necessarily derive from strictly

seasonal influences *per se*). The authors, however, use data not adjusted for seasonal varia tion. They do not, unfortunately, provide an explanation for this – a major shortcomin of the paper given that such a choice is counter-conventional. The onus, it would seem is on the authors to fully explain why they have chosen not to use seasonally adjusted data This point is further reinforced by their words:

" The annual percentage change series is somewhat smoother than the others because i removes some of the short cycles and the ARMA and Moore's filter retain some of th irregularities. However, much of the additional variance in the other filters appears to b due to seasonal cycles, and therefore these filters are limited to displaying movements i seasonally adjusted data."

In this connection, it should be noted that Moore's filter was intended to be applie to seasonally adjusted data.[3] Moreover, in a 1975 paper[4] not referenced by the authors Statistics Canada spoke specifically to the issue of seasonality within the exact context o the authors' paper: ''Obviously, a measure purporting to be the current annual rate (o the all-items CPI) can be quite misleading over the course of a year unless it is based o seasonally adjusted price indexes.''

One of the authors has informally suggested that seasonally adjusted data was not use because the seasonal adjustment process, being a smoothing one, imparts its own phas shift to the original data. When coupled with the phase shift of the smoothing applie to derive a current rate of change figure, the total phase shift of, say, the ARMA or Moore filter may not be appreciably less than the 5.5-month phase shift associated with the year over-year percentage change filter. If this is the case, then so be it – for it is clear tha any filter which retains significant seasonality simply won't do as an ''official'' measur of the current inflation rate. Even though the commonly used 12-month change has th drawback of significant phase shift, the possibility that it cannot objectively be improve upon should be approached with an open mind rather than being necessarily rejected a the outset.

It is important that a measure of the current rate of inflation be understood by the publi If it is not, it is virtually certain that the measure will never gain acceptance. It can reasonabl be argued that the use of the 12-month percentage change as the current rate of chang

wes its widespread acceptance to the fact that the algorithm is understandable and that can quickly and easily be verified from published data. On this "principle of ready verifica-on", the ARMA filter would not qualify, whereas Moore's filter and other simple mov-g average calculations may. To virtually all users, the ARMA filter would be a "black ox". It would generally not be understood nor lend itself to simple verification. It would ot see much use and it would certainly not displace the use of the 12-month percentage iange.

It can with justification be noted that the advocacy of using seasonally adjusted data tantamount to suggesting the use of another very large "black box"; it cannot be argued at the X-11-ARIMA seasonal adjustment procedure, used to produce the seasonally ad-sted CPI, lends itself to ready user verification. However, two points should be made this connection: (1) the **concept** of seasonality is generally understood even if the technical eps to eliminate it are not; and (2) the delegation of technical seasonal adjustment the central statistical agency is broadly accepted and, consequently, the resultant figures joy credibility.

The authors examine only the minimum phase-shift ARMA and Moore's filter. There e, of course, a multitude of algorithms which in principle could be contenders for the own of best measure of the current inflation rate. These contenders need to be specifically cluded in a more comprehensive review. In particular, what of the "seasonally adjusted rrent annual rate of change" (the annualized percentage change of the seasonally ad-sted all-items CPI from three months ago) that was introduced in the 1975 Statistics nada paper? It was put forward at that time as "... the appropriate measure now available " in answer to the question, "What is the **current** annual rate of change in the CPI?"[5] continued publication[6] by Statistics Canada suggests that the view held in 1975 remains rrent.

The final issue is simply that the determination of the most efficient, comprehensible easure of the current rate of inflation will still leave a major gap for the policy maker d in the expectational process. While removing seasonalities and irregularities, the filtering ocess will still leave intact impacts which may be perceived as transitory in nature - for ample, supply side shocks affecting, say, food or energy prices. While these effects are

certainly part of the current inflation rate, it may not be desirable to introdu macroeconomic policies whose impacts endure longer than the transitory influence itsel In the same vein, long-term contracts may appropriately abstract from price movemen that are perceived to be transitory in nature. Thus, the best measure of the current ra of inflation, as useful as it may be, will not sate the hunger for a measure of the ''core or ''underlying'' rate of inflation.

In order to answer better the question, ''What is the current rate of inflation?'', muc work remains undone. A choice of a base series – CPI, ISPI, implicit deflator, etc. – nee to be made, the whole issue of seasonality must be fully dealt with, the full range of poten tial filters must be evaluated, and the final result must lend itself to public acceptanc From a user's point of view, it would have been highly desirable had the authors use the 1975 Statistics Canada paper as a jumping-off point; the reader is left with questior as to how the views and analysis put forward by the authors relates to the earlier wor of Statistics Canada.

The contribution made by the paper is that it demonstrates how spectral analysis c be brought to bear on a practical and important issue of inflation measurement. It is be perceived, however, not as an answer in itself but rather as a tool in the statistician's k which can help to arbitrate among the potentially large number of sub-annual movi average processes that may be more efficient measures of the current rate of inflation th the 12-month percentage change.

Footnotes

"Sequential Signals of Recession and Recovery", Victor Zarnowitz and Geoffrey Moore, *Journal of Business*, Vol. 55, No. 1 (January 1982), p.62.

Inflation Watch, American Enterprise Institute for Public Policy Research on Behalf of the Committee to Fight Inflation, May-June 1982, pp.7-8.

ibid.

"A Note on the Current Annual Rate of Change of the Consumer Price Index", *Service Bulletin - Retail Prices and Living Costs*, Vol. 4, No. 4 (April 1975), Statistics Canada, Ottawa.

ibid.

Consumer Prices and Price Indexes, Catalogue 62-010, Quarterly, April-June 1982, p.20, Statistics Canada, Ottawa.

MESURES DU TAUX D'INFLATION ACTUEL

To provide you with a version in the official language of your choice, the French text i preceeded by the English text (p.877) in this publication.

D. Rhoades et
N. Elhawary-Rivet
Division des analyses de conjoncture
Statistique Canada

RÉSUMÉ

Les buts du texte sont de:

a) Transmettre au lecteur suffisamment de connaissances pour lui permettre une utilisa- tion optimale des taux annuels et mensuels dans l'interprétation de l'évolution d l'inflation.

b) Explorer de nouvelles façons de surveiller l'évolution cyclique de l'inflation, qui tenten d'utiliser l'information d'une façon plus efficace et par conséquent de fournir u portrait plus précis de l'économie .

On montre que la variation annuelle en pourcentage d'une série chronologique est un version lissée de la variation mensuelle en pourcentage. On montre également que la varia- tion annuelle en pourcentage retarde par rapport à la variation mensuelle, étant donn qu'elle repousse les données à 5.5 mois plus tard comparativement à la variation en pourcen tage mensuelle. Les méthodes intuitives et techniques sont présentées dans les sections I et III respectivement.

La recherche de nouvelles méthodes de mesure du taux d'inflation se préoccupe surtou de réduire le décalage inhérent au taux annuel tout en conservant le lissage qu'il accomplit On présente les propriétés de deux autres techniques. La première, proposée par Geoffre Moore, compare le niveau de prix du mois courant au niveau moyen de prix des 12 moi

récédant immédiatement celui-ci. La deuxième technique, celle de Rhoades, utilise un ltre à déphasage minimal pour lisser les séries mensuelles de variation en pourcentage. ’es méthodes ont produit des séries lisées caractérisées par un décalage chronologique lus petit que celle de la variation annuelle en pourcentage mais conservant certains cycles aisonniers qui étaient présents dans la série originale.

Pour illustrer la discussion, on emploie deux exemples pratiques utilisant l’IPVI et l’IPC. n applique aux séries chronologiques les trois méthodes et on compare les résultats en rme de lissage et de décalage chronologique.

Introduction

Dans un récent éditorial sur l’inflation, le *Financial Times of Canada* (FT,2/82) a demandé u gouvernement de “...trouver des manières de dire aux Canadiens exactement où ils en ont. Nous connaissons actuellement un taux d’inflation qui semble augmenter et diminuer n même temps. Pour que les Canadiens puissent participer à la lutte contre l’inflation, s doivent pouvoir évaluer exactement la force de l’ennemi, au lieu de se faire leurrer par es artifices statistiques”. (Notre traduction).

Essentiellement, la difficulté qu’éprouvent le *Financial Times* et, sans doute, d’autres ilisateurs vient du fait qu’ils utilisent à la fois des variations annuelles et mensuelles pour esurer le taux global d’inflation. La confusion se crée du fait qu’on peut avoir un taux ensuel qui diminue en même temps qu’un taux annuel qui progresse. La connaissance la relation qui existe entre les variations annuelles et les variations mensuelles permet concilier ces mouvements apparamment contradictoires. C’est pourquoi ce document se à offrir au lecteur une connaissance de base pour utiliser à bon escient les taux mensels et annuels lorsqu’il faut interpréter l’évolution de l’inflation.

Le présent exposé a également pour objet d’analyser de nouvelles techniques pour observer volution cyclique de l’inflation qui, par une utilisation plus efficace des données, nous nneront ainsi une image plus précise de la conjoncture actuelle. Ces autres mesures ligent toutefois l’agence statistique à mieux expliquer les chiffres qu’elle publie, et l’objtif final de la présente étude est de remplir cette obligation.

Les sections II et III, qui étudient les propriétés des variations mensuelles (un mois donn par rapport au mois précédent) et annuelles (un mois donné par rapport au même moi de l'année précédente), soulignent trois principaux points:

(i) La variation annuelle en pourcentage est une valeur lissée de la variation mensuelle

(ii) Toutes les techniques de lissage occasionnent un déphasage des signaux lissés, qu cause un retard dans l'identification des changements cycliques.

(iii) La variation annuelle en pourcentage est en retard par rapport à la variatio mensuelle, car elle décale les données de 5.5 mois comparativement à la variatio mensuelle.

La section II présente ces résultats de façon intuitive, alors que la section III le fait d façon technique. Le lecteur peut omettre cette dernière section sans perdre l'essentiel d raisonnement présenté.

La recherche de différentes mesures du taux d'inflation est axée sur la réduction du retar inhérent au taux annuel sans altération du degré de lissage.

Geoffrey Moore [2] a proposé une alternative qui compare le niveau des prix du moi en cours au niveau moyen des prix des douze mois précédents. La solution de Rhoad [4], elle, utilise un filtre à déphasage minimal pour lisser les séries de variations mensuelle. Les propriétés de ces deux solutions sont examinées à la section IV.

II Description intuitive

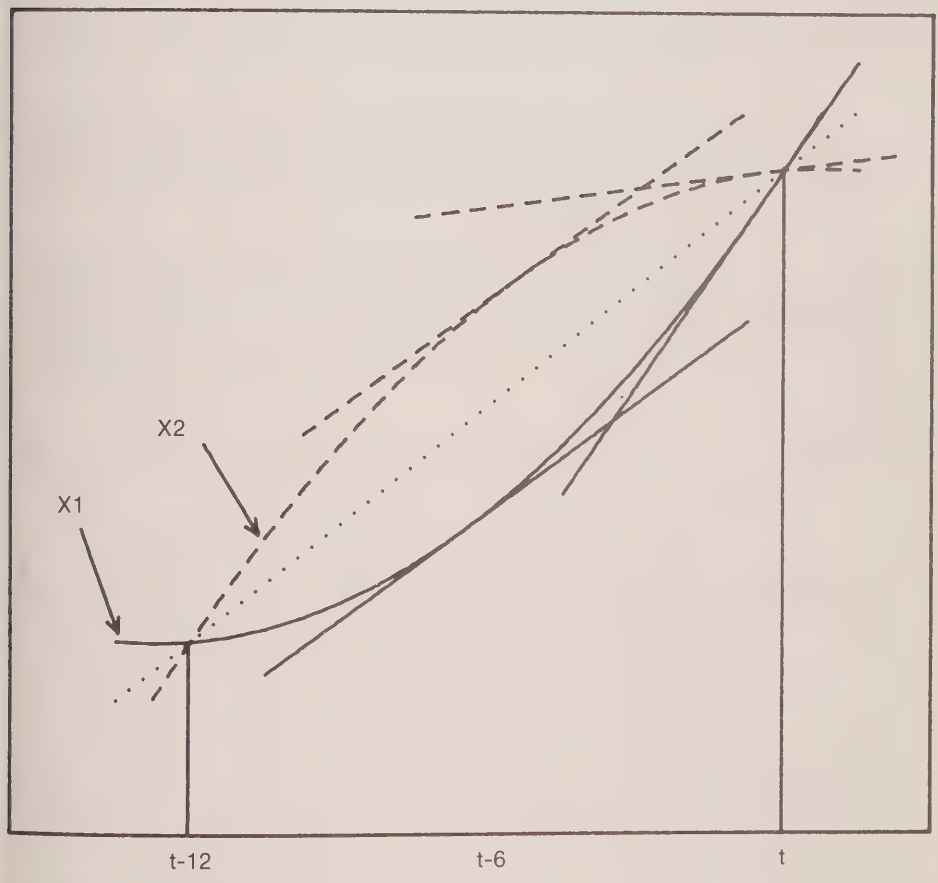
Dans la présente section, nous étudions de façon intuitive les propriétés des variatio annuelles et mensuelles en pourcentage. Le premier point que nous voulons souligner e que la variation annuelle en pourcentage est une valeur lissée de la variation mensuelle Ce point ressort clairement dans les équations suivantes:

$$\frac{x_t}{x_{t-12}} = \frac{x_t}{x_{t-1}} \cdot \frac{x_{t-1}}{x_{t-2}} \cdots\cdots \frac{x_{t-11}}{x_{t-12}}$$

$$\log \frac{x_t}{x_{t-12}} = \sum_{k=1}^{12} \log \frac{x_{t-k+1}}{x_{t-k}}$$

où on voit que la variation annuelle en pourcentage est approximé par une moyenne mobile des variations mensuelles. Prenons l'exemple suivant pour illustrer quelques autres propriétés de ces deux mesures.

Figure 1



Soit x_1 et x_2 deux séries chronologiques tracées à la figure 1 ci-dessous:

La variation annuelle en pourcentage correspond approximativement à la pente de l droite qui relie $x_1(t)$ et $x_1(t-12)$, ou $x_2(t)$ et $x_2(t-12)$. Dans le cas présent, comme $x_1(t)$ e $x_2(t)$ ainsi que $x_1(t-12)$ et $x_2(t-12)$ coïncident, la variation annuelle en pourcentage es la même pour les deux séries. Cependant, la variation mensuelle en pourcentage pour l période t est approximativement égale à la pente des droites entre t-1 et t, et elle est positiv pour x_1 mais négative pour x_2.

Deux résultats importants sont révélés à la figure 1:

(i) Comme la variation annuelle en pourcentage ne tient pas compte de la trajectoir de la courbe entre t-12 et t, elle ne peut indiquer la direction du mouvement à la date t

(ii) La pente des deux courbes à la date t-6 est approximativement égale à la variatio annuelle en pourcentage, ce qui indique que la variation annuelle représente cell apparue six mois auparavant (c'est-à-dire que cette variation a un décalage de si mois sur les séries d'origine).

Les différences entre les variations annuelles et mensuelles proviennent du fait que la variation annuelle en pourcentage dépend seulement de la valeur de x pour les périodes t et t-12 et, par conséquent:

(a) La variation annuelle ne tient pas compte des données recueillies au cours des mois entre t et t-12.

(b) La variation annuelle projette dans le présent toute particularité et irrégularité qui existaient il y a un an.

II Description théorique

Dans la présente section, nous examinons de façon plus rigoureuse les différences théori-
ues entre les variations annuelles et mensuelles en pourcentage. Pour comparer ces deux
ıesures, nous examinerons leur fonction de gain G(f) et leur fonction de phase P(f) [1]
:spectives. Les fonctions de gain et de phase de la variation mensuelle en pourcentage
voir l'annexe 1) sont:

$$G_1(f) = \left(\frac{100}{x_{t-1}}\right)^2 \cdot 4 \sin^2 \pi f \qquad (0 < f \leq .5)$$

$$P_1(f) = -.5 + \frac{1}{4f} \qquad \text{en mois} \qquad (0 < f \leq .5)$$

ı f = fréquence.

:s fonctions de gain et de phase du filtre de la variation annuelle en pourcentage sont
l: suivantes:

$$G_{12}(f) = \left(\frac{100}{x_{t-12}}\right)^2 4 \sin^2 12\pi f \qquad (0 < f \leq .5)$$

$$P_{12}(f) = -6 + \frac{1}{4f} \qquad \text{en mois} \qquad (0 < f \leq .5)$$

Il est évident que les fonctions de gain des deux filtres sont différentes,mais elles on toutes deux une forme semblable. Chaque gain a une composante sinusoïdale dont l. périodicité diffère dans chaque cas (la période est de 2 pour la variation mensuelle et d 1/6 pour la variation annuelle).

La fonction de phase présente davantage d'intérêt. La variation annuelle er pourcentage déphase la composante fréquence f d'environ six mois ultérieur (exactemen 6-(1/4f) mois), tandis que la variation mensuelle en pourcentage retarde cette composant de .5-(1/4f) mois.

La différence entre ces deux fonctions de phase représente le décalage temporel des varia tions annuelles en pourcentage par rapport aux variations mensuelles. Ce calcul indiqu que les séries de variations annuelles en pourcentage auront un retard de 5.5 mois par rap port aux séries de variations mensuelles.

IV Alternatives à la variation annuelle en pourcentage

On a montré à la section précédente que les séries de variations annuelles en pourcen tage sont une version lissée des séries de variations mensuelles et qu'elles ont un retar de cinq mois et demi par rapport à ces dernières. C'est dire qu'on doit attendre enviro six mois avant de faire une évaluation précise pour le mois en cours. L'idéal serait de trouve une alternative de lissage des séries de variations mensuelles qui aurait un plus peti déphasage. Nous examinons deux alternatives de ce genre dans la présente section.

La première, celle dont s'est servi Rhoades [4], utilise les techniques de l'analyse spec trale afin de construire des filtres qui, pour un certain degré de lissage, minimisent l déphasage.

Il s'est avéré qu'un filtre autorégressif à moyenne mobile (ARMM) avec deux coeffi cients AR et un coefficient MM[1] était approprié pour lisser les séries de variations men suelles en pourcentage. Les coefficients AR sont $b_1 = 1.451$ et $b_2 = -0.5857$, et le coeffi cient de la MM est $a_0 = 0.134$.

La deuxième alternative qui fut développée par Moore en 1982 [2] se définit comme suit

$$y_t = \frac{x_t - \frac{1}{12} \sum_{k=1}^{12} x_{t-k}}{\frac{1}{12} \sum_{k=1}^{12} x_{t-k}}$$

où y_t est la série lissée et x_t la série d'origine.

La série lissée y_t peut, dans ce cas, être aussi exprimée comme une moyenne mobile de a variation mensuelle en pourcentage, et il est donc possible d'utiliser l'analyse spectrale pour déterminer l'effet de ce filtre sur les données relatives à la variation mensuelle.

Le graphique des fonctions de gain et de phase (par rapport à la variation mensuelle n pourcentage) du filtre ARMM, du filtre de la variation annuelle en pourcentage et du iltre de Moore est présenté aux figures 2 et 3.

Pour toute fréquence donnée, la fonction de gain mesure la réduction ($G(f) < 1$) ou augmentation ($G(f) > 1$) de l'amplitude de la série filtrée, et la fonction de phase indique 'avance ou le retard de cette série. (Dans une série filtrée, l'amplitude à une fréquence lonnée est égale à l'amplitude de cette fréquence dans les données d'origine multipliée par a valeur de la fonction de gain à cette fréquence.)

Si nous considérons f = 0.04 (périodicité = 25 mois) comme la limite entre la tendance-ycle et les composantes saisonnière et irrégulière, nous pouvons voir, d'une part, que la onction de gain du filtre ARMM préserve pratiquement toute la composante tendance-ycle et, d'autre part, que les fonctions de gain du filtre des variations annuelles en pourcen-age et du filtre de Moore ne préservent pas tous les cycles. Par exemple, la variation an-uelle en pourcentage ne conservera que 68% de l'amplitude lorsque f = 0.03 (périodicité = 33 mois), ce qui montre que la variation annuelle en pourcentage réduit considérable-nent l'amplitude pour certains cycles. Par ailleurs, la fonction de gain du filtre ARMM onserve certaines des composantes irrégulière et saisonnière parce que, dans l'intervalle e fréquence [0.04, 0.2], le gain n'est pas près de zéro, ce qui signifie que les amplitudes es composantes irrégulière et saisonnière ne seront pas annulées à ces fréquences.

Figure 2
Fonction de gain

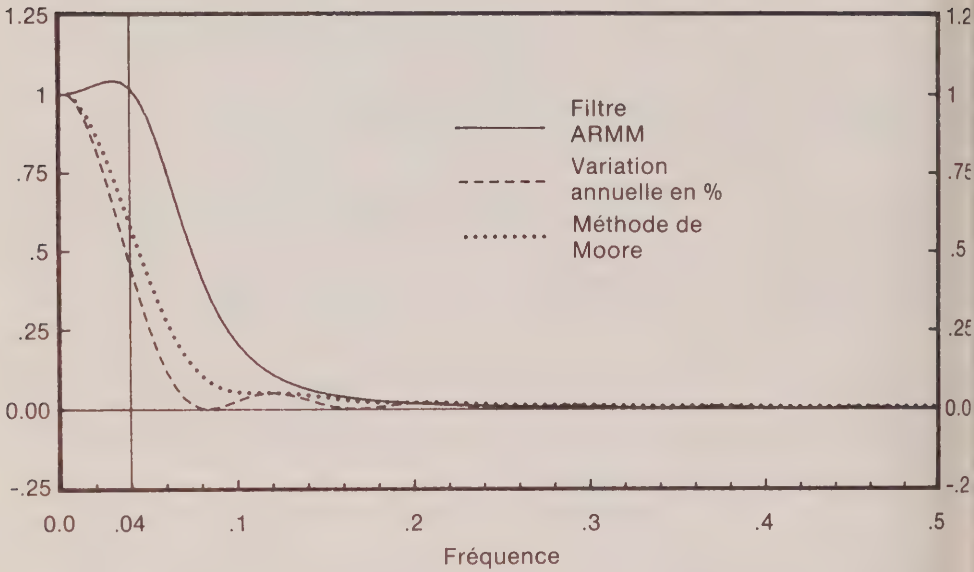


Figure 3
Fonction de phase

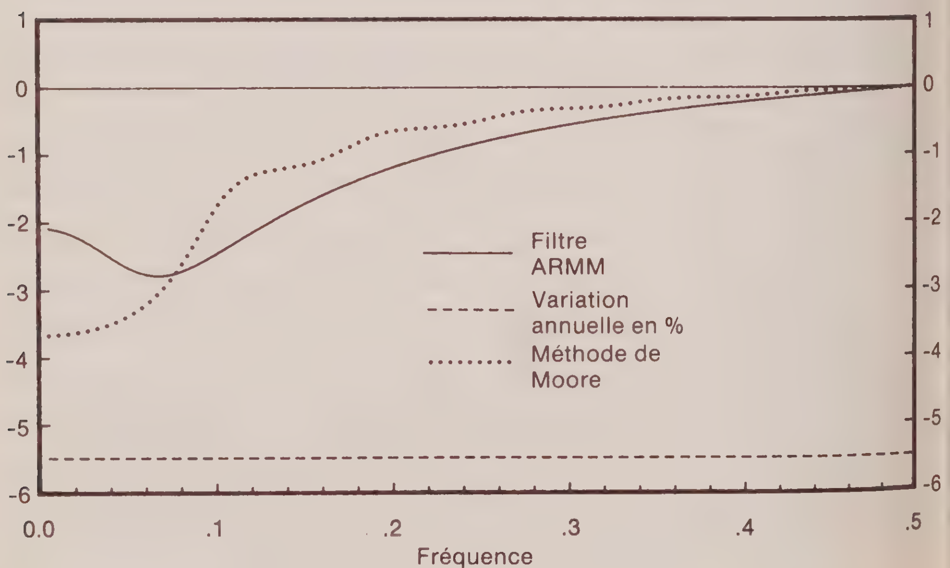


Figure 4
ISPI Percentage Change at Annual Rates
(January 1971 to September 1982)

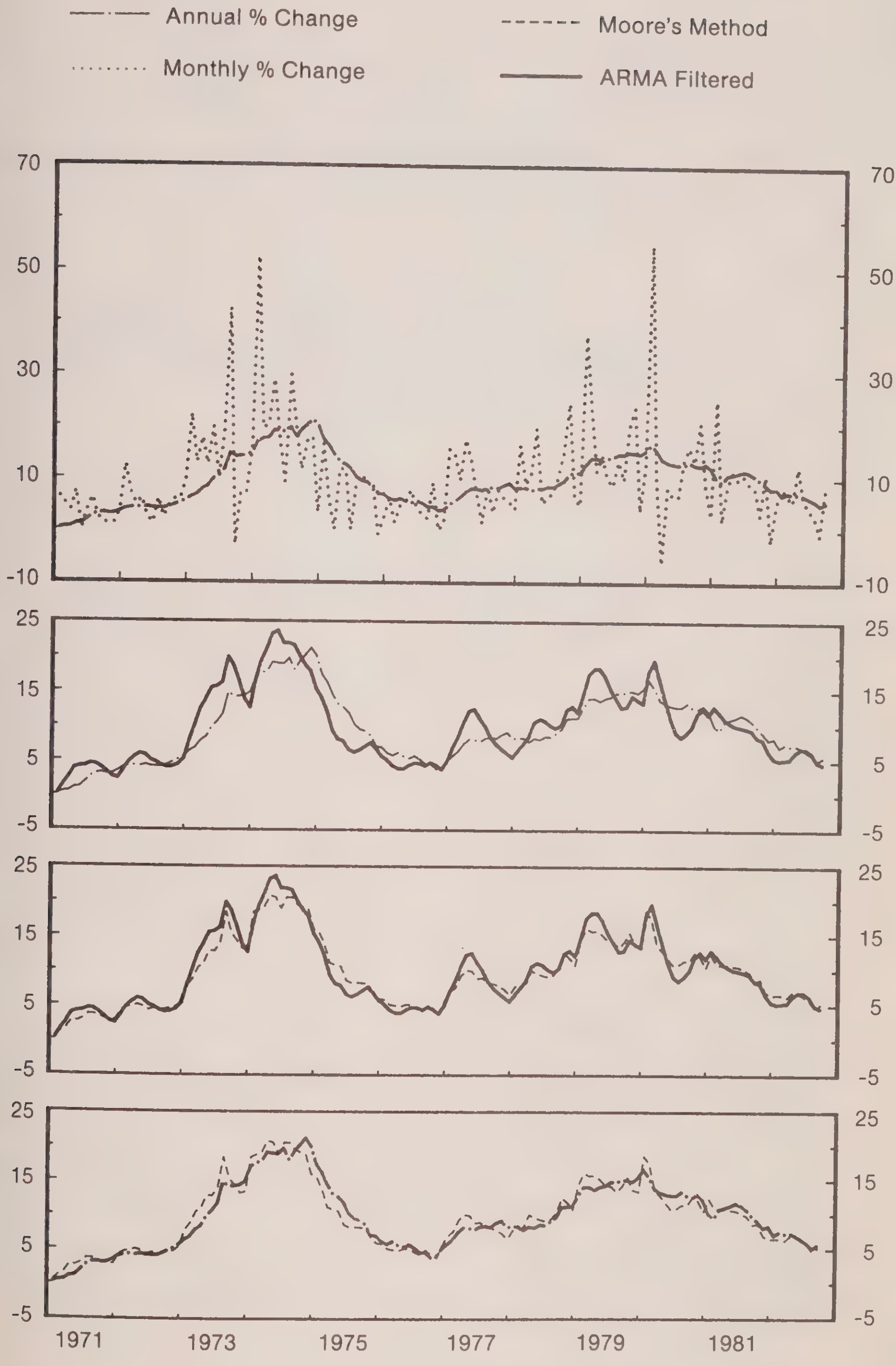
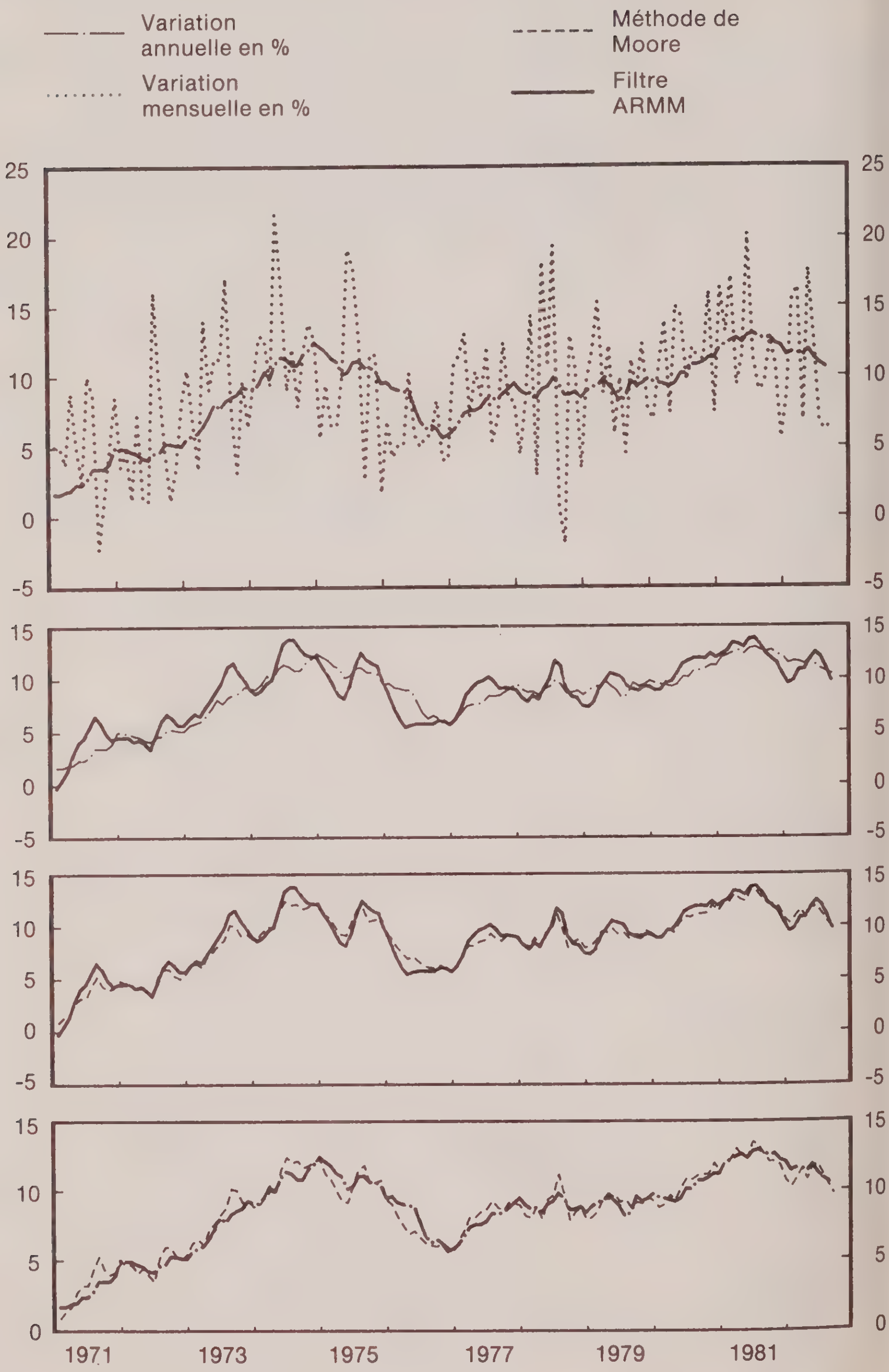


Figure 5

Variation anuelle en pourcentage de l'IPC
(Janvier 1971 à septembre 1982)



Le graphique des fonctions de phase des trois filtres mentionnés ci-dessus est présenté ı la figure 3. Le déphasage moyen du filtre ARMM pour l'intervalle de fréquence [0,0.04] ·st de deux mois environ, comparativement aux déphasages de 5.5 mois du filtre de la variaion annuelle en pourcentage et de 3.5 mois du filtre de Moore. Ainsi, malgré certaines mperfections dues au fait qu'il reproduit certains éléments des composantes saisonnière ·t irrégulière, le filtre ARMM engendre un plus petit déphasage sans altérer la composante endance-cycle. L'alternative de Moore introduit un déphasage moyen d'environ 3.5 mois t produit une courbe légèrement moins lisse que la variation annuelle en pourcentage.

.a série des variations annuelles en pourcentage de l'I.P.V.I. et de l'I.P.C. ainsi que la érie filtrée des variations mensuelles (produite à l'aide du filtre ARMM et du filtre de Aoore mentionnés ci-dessus) sont représentées aux figures 4 et 5, ce qui permet une autre omparaison des trois fonctions de gain. Il est évident que le déphasage de ces courbes st différent puisque la série des variations annuelles en pourcentage est déplacée plus à roite que la série filtrée des variations mensuelles ou que la série de Moore. La série des ariations annuelles en pourcentage est un peu plus lisse que les autres séries, étant donné u'elle supprime les composantes saisonnière et irrégulière, tandis que le filtre ARMM et : filtre de Moore reproduisent des parties des composantes irrégulière et saisonnière. outefois, une grande partie de la variance additionnelle des autres filtres semble être causée ar des cycles saisonniers, et ces filtres peuvent donc seulement révéler les mouvements : données désaisonnalisées. Le mouvement des données de Moore et de la série ajustée ar le filtre ARMM est très semblable, même dans la bande de fréquence saisonnière.

Conclusion

Nous avons montré que la variation annuelle en pourcentage est une valeur lissée de variation mensuelle en pourcentage et qu'elle est déphasée de 5.5 mois par rapport à variation mensuelle. Nous avons aussi démontré qu'il est possible de lisser la série des ıriations mensuelles en pourcentage par la méthode de Moore ou au moyen d'un filtre RMM au lieu d'utiliser le filtre de la série des variations annuelles. Les séries lissées à ıide de ces deux filtres ont un décalage moins important que la série des variations an-ıelles en pourcentage, mais elles reproduisent des cycles saisonniers qui existent dans la ·rie d'origine.

Une connaissance de ces propriétés devrait permettre aux utilisateurs de comprendre e d'interpréter les mouvements de variations annuelles et mensuelles apparemment contradic toires dans l'I.P.C. et d'autres données. Par exemple, la variation mensuelle en pourcen tage dans l'I.P.C. a chuté à un niveau de 0.7 en avril 1981, après s'être maintenu au-dessu de 1.0 durant les trois mois précédents. Cette diminution avait été prévue par plusieur analystes alors que les prix industriels et des matières premières avaient déjà ralentis en réac tion à la récession du premier semestre de 1980. En rétrospective, avril semble avoir ét le début d'un relâchement des taux d'inflation qui avaient persisté jusqu'à ce moment-là Malheureusement, la variation annuelle en pourcentage est passée à 12.6 en avril par rap port à 12.4 le mois précédent, et ce fait a porté à confusion la perception des événement de certaines gens à cette date. Cette augmentation de la variation annuelle à un niveau de 12.6 fit la manchette dans les journaux alors que ce taux éclipsait de 0.1% le précéden taux record d'après-guerre. Donc, au moment où l'inflation avait commencé à diminuer plusieurs personnes pensaient qu'elle augmentait pour atteindre un nouveau sommet, c qui était attribuable au fait que ces personnes utilisaient une mesure qui était périmée d six mois. Afin de réduire la possibilité de telles interprétations, il est recommandé que tan et aussi longtemps qu'on utilisera les variations mensuelles et annuelles en pourcentage les communiqués de presse comprennent une note explicative reflétant les propriétés relative au lissage et à l'actualité de ces variations. On devrait toutefois poursuivre la recherch d'une autre méthode et adopter la plus efficace.

Annexe 1 Fonction de transfert des filtres de variations annuelles et mensuelles en pourcentage

La variation mensuelle en pourcentage

$$y_t = \frac{x_t - x_{t-1}}{x_{t-1}} \; 100$$

peut être considérée comme une moyenne mobile qui varie en fonction du temps:

$$y_t = a_0 x_t + a_1 x_{t-1} \quad \text{où} \quad a_0 = \frac{100}{x_{t-1}} = -a_1$$

La fonction de transfert [1] de cette moyenne mobile est définie par:

$$A_1(f) = a_0 + a_1 e^{-2\pi i f} = \frac{100}{x_{t-1}} (1 - e^{-2\pi i f})$$

$$= \frac{100}{x_{t-1}} e^{-\pi i f} (e^{\pi i f} - e^{-\pi i f})$$

$$= \frac{100}{x_{t-1}} e^{-\pi i f} (2i \sin \pi f)$$

$$= \frac{100}{x_{t-1}} e^{-\pi i f} \cdot e^{\frac{\pi}{2} i} (2 \sin \pi f) \quad \text{où } i = e^{\frac{\pi}{2} i}$$

$$= e^{(\frac{\pi}{2} - \pi f) i} \; \frac{200}{x_{t-1}} \sin \pi f$$

La fonction de gain $G_1(f)$ est le module quadratique $A_1(f)^2$ de la fonction de transfert, et la fonction de phase exprimée en radians, $P_1(f)$, est l'angle complexe de $A_1(f)$. Donc, pour la variation mensuelle en pourcentage:

$$G_1(f) = 4\left(\frac{100}{x_{t-1}}\right)^2 \sin^2 \pi f$$

$$P_1(f) = \frac{\pi}{2} - \pi f$$ en radians

$$\frac{P_1(f)}{2\pi f} = \frac{1}{4f} - .5$$ définit le déphasage en périodes.

Il est intéressant de noter que la fonction de gain est la seule qui varie dans le temps.

Un raisonnement analogue conduit aux fonctions de gain et de phase des variations annuelles en pourcentage.

$$y_t = a_0 x_t + a_{12} x_{t-12} \quad \text{où } a_0 = \frac{100}{x_{t-12}} = -a_{12}$$

$$A_{12}(f) = a_0 + a_{12}\, e^{-24\pi i f}$$

$$= e^{(\frac{\pi}{2} - 12\,\pi f)i} \left(\frac{200}{x_{t-12}} \sin 12\pi f\right)$$

$$G_{12}(f) = 4\left(\frac{100}{x_{t-12}}\right)^2 \sin^2 12\pi f$$

$$P_{12}(f) = \frac{\pi}{2} - 12\pi f$$ en radians

$$\frac{P_{12}(f)}{2\pi f} = \frac{1}{4f} - 6$$ en périodes.

Tel qu'il a été mentionné à la section II, on peut considérer la variation annuelle en pourcentage comme une valeur lissée de la variation mensuelle. $A_1(f)$ et $A_{12}(f)$ fournissent des

renseignements sur le comportement des variations annuelles et mensuelles en pourcentage par rapport aux données d'origine x. Il est aussi intéressant de comparer le comportement des variations annuelles et celui des variations mensuelles en pourcentage. On peut démontrer [1, p.45] que la fonction de transfert qui exprime la relation entre les variations mensuelles en pourcentage et les variations annuelles est définie par l'équation $A(f) = A_{12}(f) / A_1(f)$.

$$\text{Alors } G(f) = \frac{G_{12}(f)}{G_1(f)} = \left(\frac{x_{t-1}}{x_{t-12}}\right)^2 \left(\frac{\sin 12\pi f}{\sin \pi f}\right)^2$$

$$\text{et} \qquad P(f) = P_{12}(f) - P_1(f) = -11\pi f \quad \text{en radians}$$

$$\text{ou} \qquad \frac{P(f)}{2\pi f} = -5.5 \quad \text{est le déphasage exprimé en mois.}$$

Ces résultats montrent que la variation annuelle en pourcentage est une valeur lissée de la variation mensuelle et qu'elle est constamment déphasée de 5.5 mois.

Renvoi

1 Le filtre prend la forme de $y_t = b_1 y_{t-1} + b_2 y_{t-2} + a_0 x_t$ où x_t sont les données d'origine et y_t sont les données lissées.

Référence

[1] Jenkins, G.M. et D.G. watts, *Spectral Analysis and Its Applications,* Holden-Day, 1969.

[2] *Inflation Watch,* Mai-juin 1982, American Enterprise Institute for Public Policy Research.

[3] Nerlove, M., *Spectral Analysis of Seasonal Adjustment Procedures,* Econometrica, Vol. 32, juillet 1964.

[4] Rhoades, D., *La conversion de l'actualité en fiabilité des séries chronologiques économiques,* Revue canadienne de statistique, février 1980.

'HE ESTIMATION OF SEASONAL VARIATIONS IN CONSUMER PRICE INDEXES

'our vous fournir une version dans la langue officielle de votre choix, le texte anglais est uivi du texte français (p.968) dans cette publication.

:stela Bee Dagum and
Aarietta Morry
'ime Series Research and Analysis Division
'tatistics Canada

'UMMARY

This study analyses the characteristics of seasonal variations in the All-items Consumer 'rice Index and its major two subcomponent series, Food and All-items-excluding Food. ' evaluates which of the two traditional seasonal adjustment procedures, direct or indirect, · to be preferred from the viewpoint of the degree of smoothness of the monthly rate f growth. Other related problems discussed are the application of the X-11-ARIMA :asonal adjustment method with or without the ARIMA extrapolation option and con- ırrent versus year-ahead factors. The selection of the optimal procedure is made accor- ing to the degree of smoothness and size of the revisions of the seasonally adjusted series.

. Introduction

An important use of seasonally adjusted consumer price indexes is the derivation of 'owth rates in these series. Target rates of growth may be set by the government as a ıide to policy, and actual rates of growth, excluding seasonal variations, would then be ›mpared with these targets. How close the actual rates come to the target rates depends part on the quality of the seasonal adjustment. Errors in the seasonal adjustment pro- ıce errors in the seasonally adjusted rates of growth and lead to uncertainty about how osely the targets are being met. Therefore, revisions of current seasonally adjusted series ıve always concerned policy makers, particularly if the revisions are high or change the rection of the cyclical movements.

Another desirable property to policy makers is the smoothness of the seasonally adjuste data. It is expected that removal of the seasonal component eliminates a source of recurr ing variability in the series, and thus that the seasonally adjusted series be less rough tha the original.

This study uses these two criteria, degree of smoothness and size of the revisions, t decide which is the optimal seasonal adjustment procedure for the All-items Consume Price Index and its major two subcomponents.

Section 2 introduces the main characteristics of seasonal variations in the consumer pric index series. Section 3 discusses different seasonal adjustment procedures, regression ver-sus moving average models, concurrent versus year-ahead seasonal factors and direct ver-sus indirect seasonal adjustment of aggregated series. Section 4 gives a set of measure for the two evaluation criteria chosen in this study, the degree of smoothness and th magnitude of the revisions of the seasonally adjusted series. Section 5 analyses the em-pirical results for the All-items CPI, and its two subcomponents, Food and All-items excluding Food. Finally, the main conclusions of this study are given in Section 6.

2. Seasonal Variations in the Consumer Price Indexes

In the analysis of monthly consumer price indexes it is very important to estimate th impact of seasonal variations. Most of the change in month-to-month comparisons ma be due to seasonality instead of other underlying movements such as trend and cycle.

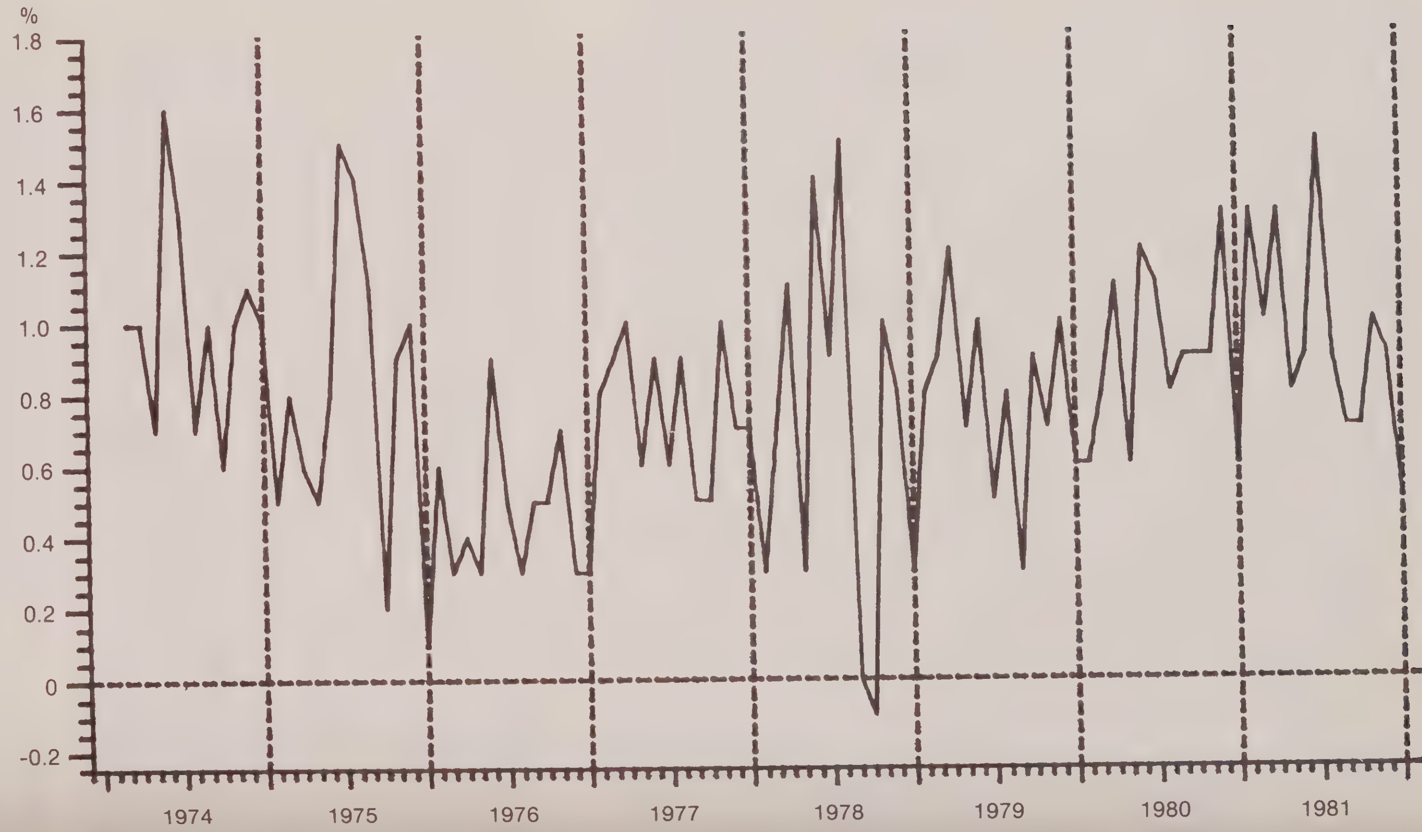
Seasonal variations represent the composite effect of climatic and institutional event which repeat more or less regularly each year. In the particular case of consumer pric indexes, seasonality mainly results from two causes, the climate, (which affects the de-mand and the supply of commodities) and the frequency of price collection which depend on the nature of the commodity. Besides the climate other factors can affect seasonalit of purchases of some commodities. On the supply side, the evolution of technology an the expansion of foreign trade may reduce the impact of the weather on the availabilit of the goods and services. On the demand side, new consumer preferences and the leve of income of the population may change the seasonal behavioural pattern of consumers Whether due to varying supply or demand, a seasonal decline in the purchase of a seasona

commodity (e.g., strawberries, corn, winter coats, bicycles) is sometimes severe enough hat it is not possible to have a representative price sample for it in the out-of-season period. f weights are fixed for all the months of the year, price changes must be imputed to seasonal commodities during the out-of-season period. The practice at Statistics Canada is to carry he last in-season price available through the out-of-season months. Thus, the month-to-nonth change for the first in-season month actually represents the price change between hat month and the last month of the previous in-season period. This method has the ad-antage of relying exclusively on observed prices while preserving the fixed-basket con-ept. It has the disadvantage, however, when there is an upward trend in prices, of nderestimating the aggregate month-to-month price movement in the out-of-season period nd overestimating it for the first in-season month. The other main factor that introduces easonality "but" in commodities which are not really seasonal, is the frequency of price ollection. The frequency of price collection depends on the nature of a given commodity. ioods and services subject to frequent price change require more frequent price collec-on. After April 1982, prices for food items are collected twice a month. Most other com-nodities are monthly priced, namely, household supplies, household furnishings, clothing, asoline, pharmaceuticals, personal care supplies, tobacco products, alcoholic beverages, ent, mortgage interest rates and new houses.

The remaining commodities, however, have prices collected at intervals larger than one onth and the current practice is to repeat the last available price until a new one is col-cted. Prices for household appliances are collected six times a year. Prices for automobiles, othing services and personal care services are collected quarterly; prices for newspapers e recorded twice a year; automobile registration fees and property taxes are recorded nually. Furthermore, additional price collections are carried out for these commodities hen there is evidence of a significant price change between regular pricing dates (see *The onsumer Price Index Reference Paper, Concepts and Procedures,* 1982). This varying equency of price collection introduces seasonal variations which follow patterns significant-different from those of monthly priced commodities.

Generally, one can get a good idea of the type of seasonality present by looking at the ries. However, in the case of the month-to-month percentage changes in the all-items PI (Figure 2.1) it is difficult to infer whether seasonality is present, therefore we resort the technique of spectral analysis.

Figure 2.1
Unadjusted C.P.I. (All Items) Month to Month Percentage Change



Information on the main characteristics of the seasonal variations affecting a given series :an be drawn from the spectrum of the series. Although the spectrum is strictly based on he stationarity assumptions and most of economic time series are not stationary, this pro- blem can be removed in the case of actual data analysis with the concept of the pseudo- pectrum (Hatanaka and Suzuki, [1967]). The pseudo-spectrum is essentially the spectrum stimated by the computer as though the data were stationary. It can be thought of as he average of a time-changing spectrum.

If the seasonality is perfectly periodic, in the sense that $S_t = S_{t-12}$, then it can be epresented in the frequency domain by the following simple model,

$$S_t = \sum_{j=1}^{6} (\alpha_j \cos 2\pi\lambda_j t + \beta_j \sin 2\pi\lambda_j t); \; \lambda_j = j/12; \; \beta_6 \equiv 0 \qquad (2.1)$$

/here the λ_j $j = 1,2,3,\ldots,6$ are the seasonal frequencies and the corresponding periods of ieir cycles (in months) are equal 12, 6, 4, 3, 2.4 and 2 respectively.

The model (2.1) is said to be deterministic if α_j and β_j are constants and stochastic if $_j$ and β_j are purely random mutually uncorrelated variables.

If a series follows model (2.1) its spectrum will be zero except for six spectral lines at e seasonal frequencies λ_j. For most economic time series, however, seasonality is not perfectly periodic process but it changes gradually.

In such cases, the model (2.1) may be written as,

$$S_t = \sum_{j=1}^{6} (\alpha_{jt} \cos 2\pi\lambda_j t + \beta_{jt} \sin 2\pi\lambda_j t); \; \lambda_j = j/12; \; \beta_{6t} \equiv 0 \qquad (2.2)$$

Where α_{jt} and β_{jt} are slowly varying parameters. α_{jt} and β_{jt} can either be deterministic nctions of time or can be stochastic processes with spectra dominated by low frequency mponents.

Series that follow model (2.2) have spectra characterized by peaks at the seasonal frequencies λ_j. Their spectra will be zero outside the seasonal frequency bands $\lambda_s(\delta)$ defined by,

$$\lambda_s(\delta) = \left\{\lambda_s \text{ in } (\lambda_j - \delta, \lambda_j + \delta),\ j = 1,2,\ldots,5,(\lambda_6 - \delta, \pi)\right\} \qquad (2.3)$$

Where the value of δ depends on the rate at which α_{jt} and β_{jt} change in value.

The narrower the peaks, the more regular will be the seasonal variations. If the seasonal variations are estimated by moving average methods, the shorter the moving average, the faster the seasonal estimate may vary and the broader the band of frequencies about each seasonal frequency is supposed to be.

The power spectrum at those frequencies lower than $\lambda = \frac{1}{18}$ are generally attributed to business cycle variations and the power at the remaining frequencies which fall outside the seasonal bands, to the irregular fluctuations.

The spectra of the three series analysed here have been estimated after removing the power of an exponential trend.

Looking at the estimated spectrum of the All-items CPI in Figure 2.2, it is interesting to observe that there is no power at the fundamental seasonal frequency $\lambda_1 = \frac{1}{12}$ and that the dominant seasonal peak occurs at $\lambda_6 = \frac{1}{2}$ corresponding to a two-month cycle. The other two most important peaks occur at λ_5 and λ_3 which correspond to 2.4-and 4-month cycles, respectively. The majority of economic time series are characterized by an opposite type of spectrum with higher peaks at lower seasonal bands, around the fundamental and its first harmonic (see e.g. Nerlove, Grether and Carvalho, [1979]).

Another important feature is that these seasonal bands are broad, indicating that seasonality changes fast.

The All-items CPI is the aggregate of 490 elementary commodity groups. These elementary commodity groups can be aggregated to form two major subgroups, Food and All-items-excluding Food, which have different seasonal causes as discussed above. We also estimated the spectra of these two series.

consommation (tous articles) — Pourcentages de variation mensuelle

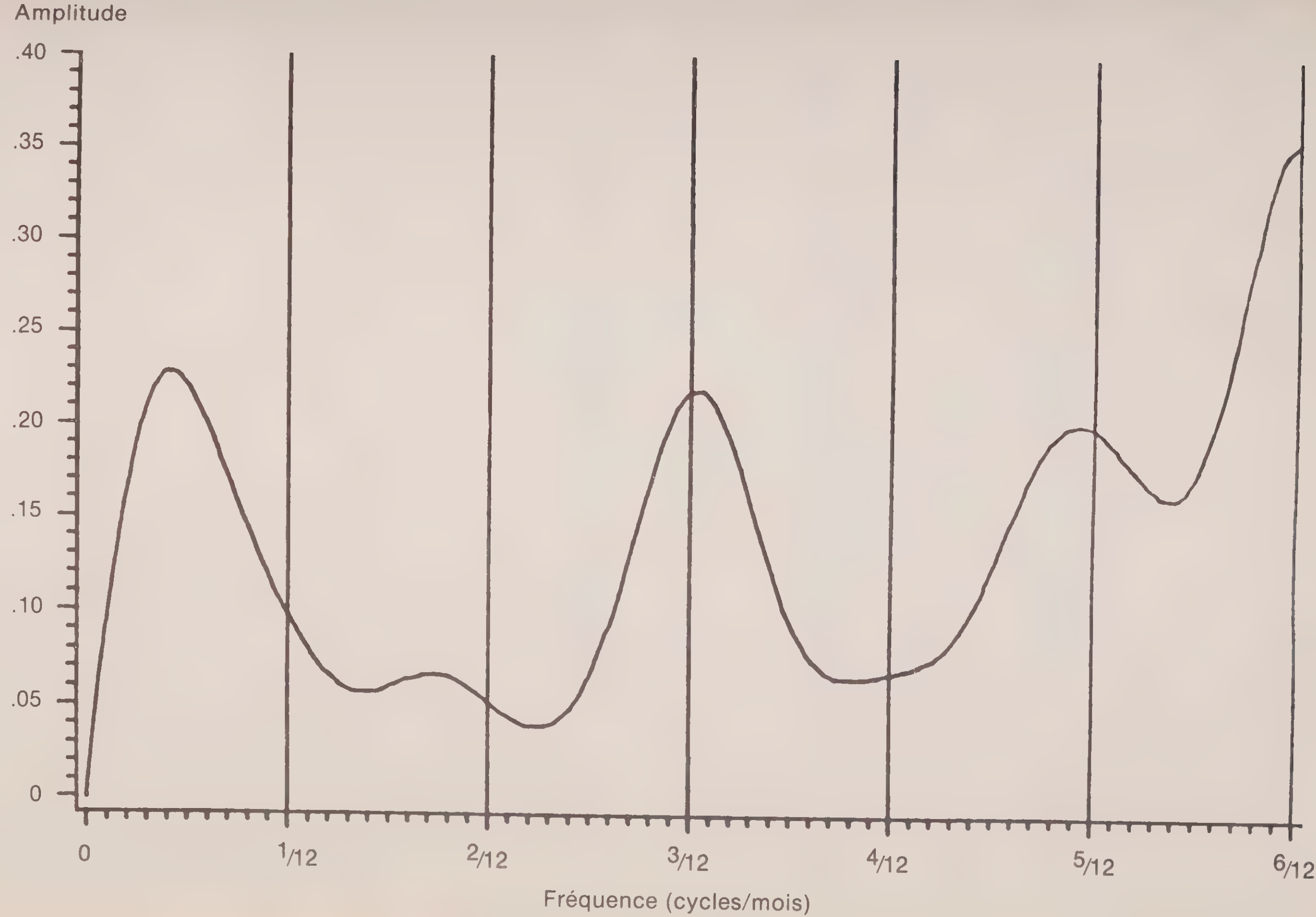
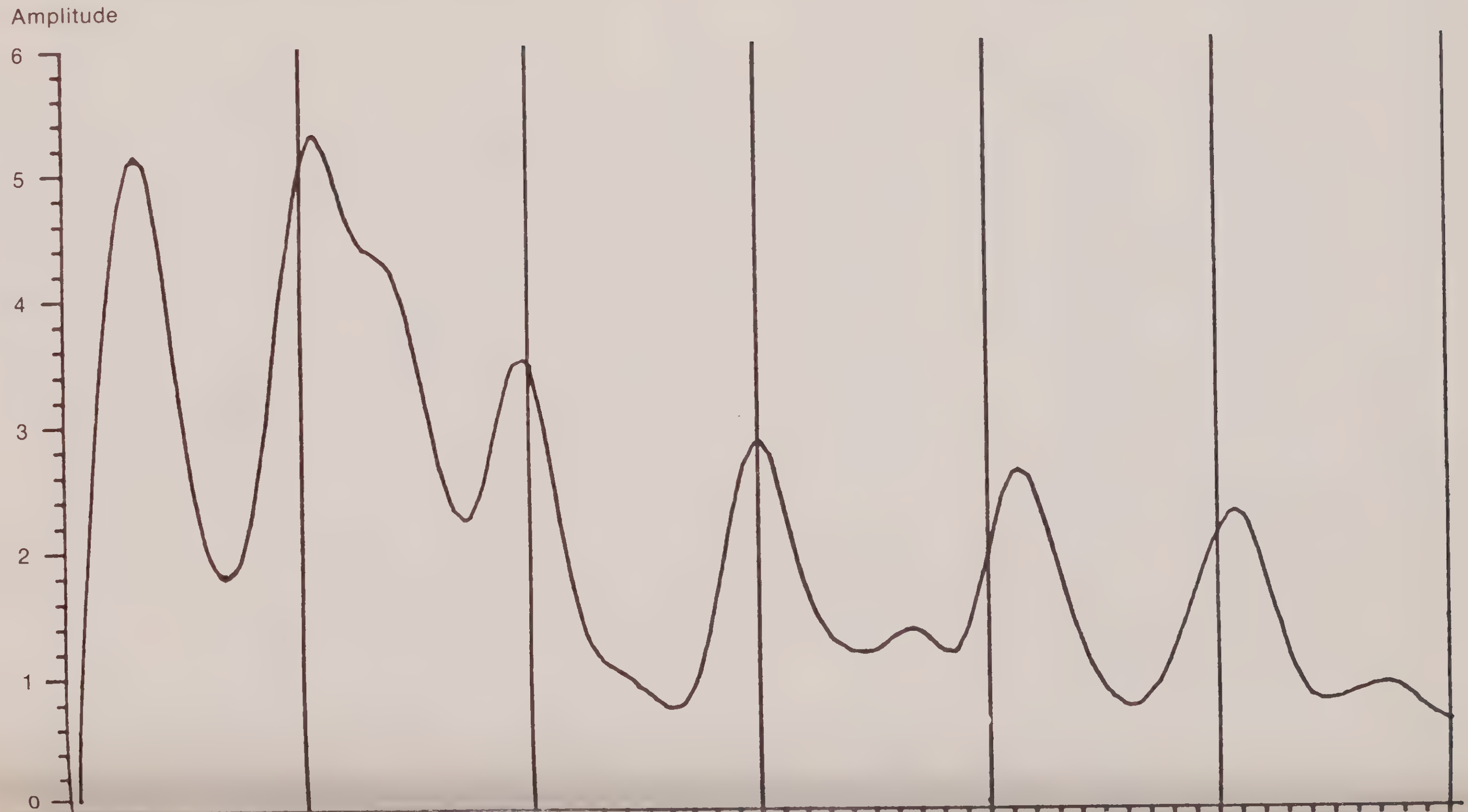


Figure 2.3
Power Spectrum Estimates of Unadjusted C.P.I. (Food) Month to Month Percentage Change



(All Items Excl. Food) Month to Month Percentage Change

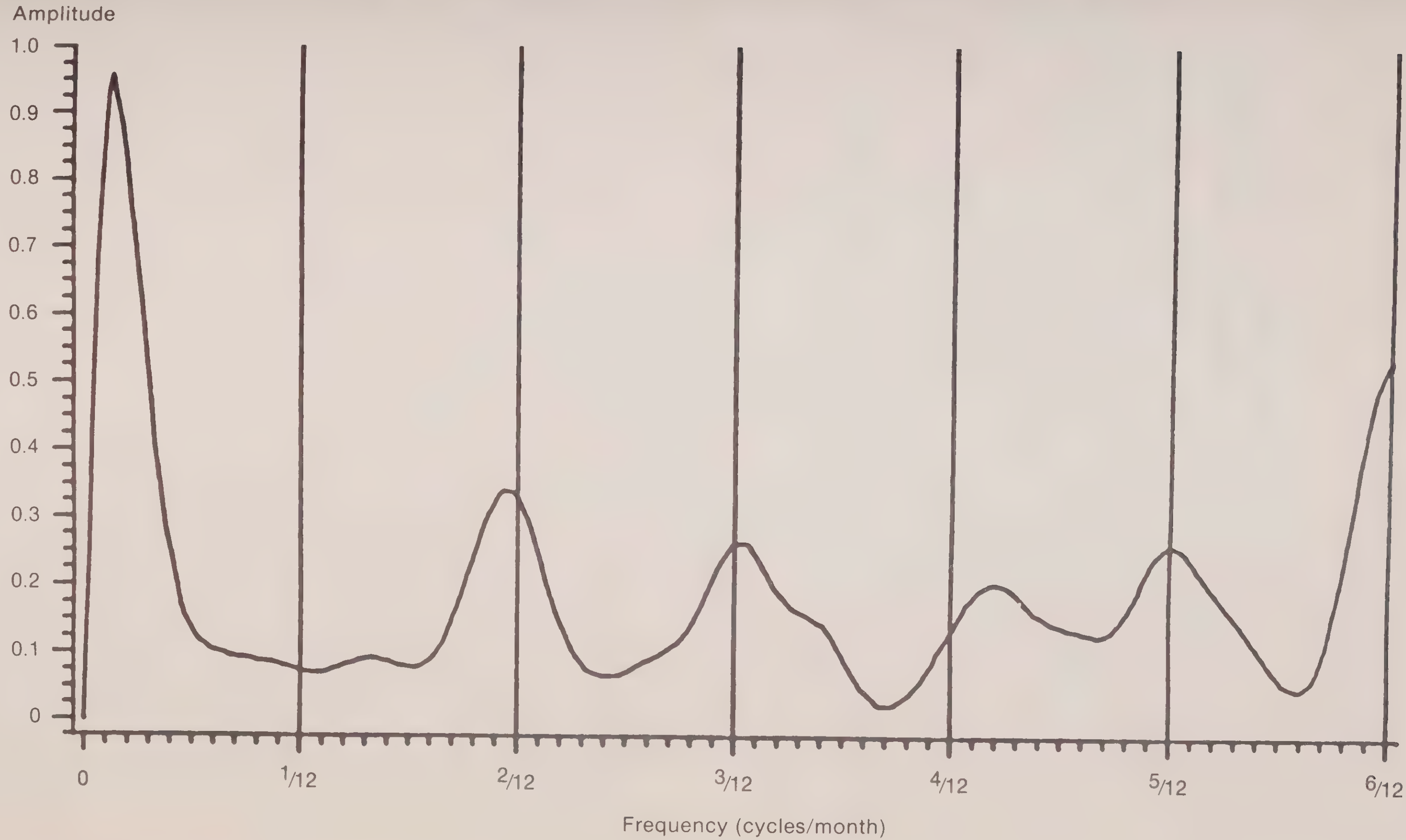


Figure 2.3 shows the spectrum of the Food Price Index which looks very similar to tha of main economic indicators.

It also should be noted that there is a high power at the low frequencies correspondin to three- and four-year cycles probably caused by the beef and hog cycles.

Figure 2.4 shows the spectrum of All-items-excluding Food Price Index. Contrary tc the previous case, the power of the spectrum is higher at the higher seasonal frequencies Two peaks are shown at λ_2 and λ_6 corresponding to the six- and two-month cycles, anc high power is also observed at λ_3, λ_4 and λ_5 whereas there is none at the fundamenta seasonal frequency band. Many items included in this series are priced at different period longer than one month and, in fact, this varying frequency of price collection introduce the artificial type of seasonality revealed by the series spectrum. Furthermore, the seasona bands are broader than those observed in the Food Price Index Series implying a rapidl changing seasonality.

3. Seasonal Adjustment of Consumer Price Indexes

The seasonal adjustment of economic time series generally poses the problems of decidin on the method to be applied and whether a concurrent or a forecast seasonal factor shoul be used for the current observation. Furthermore, if the series in question is an aggregat as the consumer price index, a decision must be made concerning the level of aggregatio at which the seasonal adjustment should be done.

The selection of the ''best'' procedure for each situation, depends on a set of criteri which are determined mainly by the purpose of the seasonal adjustment. In the case o the consumer price index the criteria chosen here are the size of the revisions and the degre of smoothness of the seasonally adjusted data.

3.1 Seasonal Adjustment Methods

The majority of seasonal adjustment methods developed thus far are based on univariat time series models where the estimation of the seasonals is made in a simple and mechanica manner and not based on a causal explanation of the phenomenon under study. Very few

ttempts have been made to follow the latter approach and none of them reached further ıan the experimental stage.

On the other hand, univariate time series methods of seasonal adjustment try to estimate ıe generating mechanism of the observations under the simple assumption that the series composed of a systematic part and a random part. The feasibility of this decomposition as proved by Wold [1938] with a theorem stating that if a time series X_t is stationary the second order (i.e. its mean does not depend on time t and its autocovariance depends ıly on the time lag), then it can by uniquely represented as the sum of two mutually un- ›rrelated processes, one a linearly deterministic process ξ_t and the other an infinite mov- g average process η_t. Hence,

$$X_t = \xi_t + \eta_t; \ \eta_t = \sum_{j=0}^{\infty} \beta_j U_{t-j}; \ \Sigma\beta_j^2 < \infty \tag{3.2}$$

ıere

$$\begin{aligned} E(U_t U_s) &= \sigma^2 && t = s \\ &= 0 && t \neq s \\ E(\eta_t \xi_t) &= 0 && \text{for all } t,s \end{aligned}$$

ıd ξ_t is linearly deterministic. This decomposition is still valid for non-stationary series ːh as economic time series if they can be made stationary by suitable transformations.

Within the large class of univariate time series models, two main categories can be tinguished, the regression models and the moving average models. A brief description the main assumptions implied by both kinds of models follows.

Most work in regression models for seasonal adjustment is based on the assumption t the systematic part of a time series can be approximated closely by simple functions time over the **entire span** of the series.

In general, two types of functions of time are considered. One is a polynomial of fair low degree that fulfills the assumption that the economic phenomenon moves slowl smoothly and progressively through time (the trend). The other is a linear combinati of sines and cosines of different frequencies representing oscillations, strictly periodic not, that affect also the total variation of the series (the cycle and the seasonals).

These functions are estimated by least squares methods. To obtain efficient estimate the random component must be independent; otherwise, an appropriate version of general ed least squares must be applied. If the relationship among the components (trend, cyc seasonality, irregulars) is assumed to be multiplicative, the standard procedure is to ta logarithms and their differences to transform the generating mechanism of the series in an additive form with a stationary random part.

Major contributions to the development of this kind of regression model were ma by Hannan [1960], Lovell [1963], Rosenblatt [1963], Ladd [1964], Jorgenson [1964] a Henshaw [1966].

To overcome the limitations of using a global representation of the trend-cycle, Duv [1966] and Stephenson and Farr [1972] used local polynomials (spline functions) for su cessive short segments of series. These regression models, however, implied a determinis behaviour of the time series components. Later, Pierce [1978], Gersovitz and MacKinn [1978] and Havenner and Swamy [1978] took into consideration the possibility of stocha behaviour of the trend-cycle and seasonal components by developing mixed models or regr sion models with time varying parameters.

Regression models have been seldom used by official statistical agencies for the seaso adjustment of their data. The main reasons for this being the fact that until very recen the methods developed assumed a deterministic behaviour of the components and, th have not yet overcome the problem that the seasonally adjusted series are wholly revi when new observations are added; recent estimated values are strongly influenced by tant observations.

The majority of the seasonal adjustment methods applied by statistical bureaus belong the class of moving average models which make the assumption that although the signal a time series is a smooth function, it cannot be approximated well by simple mathematical ınctions over the entire range.

Since moving averages are linear transformations they have the properties of scale preser-ıtion and additivity. Furthermore, they have the time invariance property which is not ıared by the regression methods. The time invariance property means that if two inputs, $_t$ and $X_{t+\tau}$, to the moving average L, are the same except for the time displacement ıen the outputs LX_t and $LX_{t+\tau}$ are also the same except for their time displacement. other words, the moving average or linear filter L responds always in the same manner.

Methods based on moving averages techniques assume that the trend-cycle and seasonal ımponents change through time in a stochastic manner. The seasonal adjustment methods at belong to this category are mainly descriptive non-parametric procedures in the sense at they lack explicit parametric models for each unobserved component.

In most recent years, however, several attempts have been made to develop model-based ocedures where univariate statistical models are explicitly assumed for each component. ıese are not causal or explanatory models, and thus, the stability of their estimated ırameters is seriously jeopardized when the series are affected by exogeneous events.

The explicit models mainly belong to the Gaussian ARIMA (autoregressive integrated oving average) type developed by Box and Jenkins [1970] or to variations of it (see e.g. ırman, [1980]; Hillmer and Tiao, [1982] and Nerlove, Grether and Carvalho, [1979]). her types of models (not ARIMA) have been assumed by Akaike and Ishiguro [1980] d Bilongo and Carbone [1981].

These new methods are still in a developmental stage and the majority of the moving erage procedures officially adopted by statistical bureaus belong to the non-parametric pe (see Kuiper, [1978]). Among these, the Method II - X-11 variant developed by Shiskin, oung and Musgrave [1967] and the X-11-ARIMA developed by Dagum [1975 and 1980] e the most widely applied. The X-11-ARIMA is a modified version of the X-11 variant ıich basically consists of extending the original series, at each end, with extrapolated

values from ARIMA models, and then, seasonally adjusting the extended series. The ne set of moving average weights results from the combination of the X-11 seasonal filte with the extrapolation ARIMA filters. The seasonal adjustment filters of X-11 ar X-11-ARIMA differ both as applied to the last available observation and for the data the most recent years. Only the symmetric filter applied to central observations is the san for both procedures. If the ARIMA option is not applied, the X-11-ARIMA reduces the X-11 method.

The revisions of the current seasonally adjusted values by X-11-ARIMA with or witho the ARIMA option (X-11) as well as for any method based on moving average model are due to: (1) differences in the moving averages or smoothing linear filters applied the same observation as later data become available; and (2) the innovations that ent into the series with new observations.

In contrast with seasonal adjustment regression methods, revisions of the seasonally a justed data stop when the symmetric filters are applied. In the context of the two metho analysed here, this means that the first and last three and a half years of a given seri will be revised because a good approximation of their filter requires seven years of da to produce a central estimate (see Young, [1968] and Wallis, [1974]).

The total revision due to differences between the non-central and central filters appl to the same observation as it changes its position relative to the end of the series has be measured for both methods by one of the authors (Dagum, [1982a and 1982b]). The resu obtained clearly indicated that the use of extrapolated values from ARIMA models reduc significantly the filter revision measures. The revision measures were much smaller at t lower seasonal frequency bands (mainly around the fundamental seasonal frequency) th at the higher ones.

The conclusions from this theoretical study conform to the results given in several e pirical works (Dagum [1975 and 1978]; Farley and Zeller [1978]; Kuiper [1978 and 198 Pierce [1980] and Kenny and Durbin [1982]).

Under the assumption that outliers have been removed and the series is additively deco posed, the total revisions of real seasonally adjusted series depend on the size of the fil

evision times the power spectrum of the series; the total revisions will be larger the larger he power spectrum of the input series in the seasonal frequencies. Therefore, series with mall seasonal variations or with very regular seasonal patterns may be equally well seasonal- adjusted by X-11-ARIMA with or without the extrapolated values.

The seasonal adjustment method adopted by Statistics Canada is the X-11-ARIMA. lowever, because the consumer price index series analysed in this study have seasonal varia- ons with special characteristics, tests have been made to assess whether the use of the RIMA option improves or not the accuracy of the current seasonally adjusted series.

.2 Concurrent Versus Forecast Seasonal Factors

The information given by seasonally adjusted series plays a crucial role in the analysis f current economic conditions, particularly, in determining the stage of the cycle at which e economy stands. Such knowledge is useful in forecasting subsequent cyclical movements nd provides the basis for decision-making to control the level of economic activity.

An important use of seasonally adjusted consumer price indexes is the derivation of owth rates in these series. Target rates of growth may be set by the government as a ide to policy, and actual rates of growth, excluding seasonal variations, would then be mpared with these targets. How close the actual rates come to the target rates depends part on the quality of the seasonal adjustment. Errors in the seasonal adjustment pro- ce errors in the seasonally adjusted rates of growth and lead to uncertainty about how sely the targets are being met. Therefore, revisions of current seasonally adjusted series ve always concerned policy makers, particularly if the revisions are high or change the ection of the cyclical movements.

A **current** seasonally adjusted value can be obtained by either applying a seasonal factor ecast or a concurrent seasonal factor. The latter is generated by seasonally adjusting, h month, all the data available up to and including that month. The former is obtained m the seasonal adjustment of a series that ended one year before.

he current practice for the consumer price indexes series is to apply the seasonal ad- ment method once a year to data ending in December of that year (year t) and the

forecast or projected factors for the year ahead (year $t+1$) are then used to seasonally adjust the data as they become available.

The use of forecast seasonal factors has usually been justified on the basis of increasing public confidence in the seasonally adjusted key economic indicators since these factors are released ahead of the months to which they are applied. However, the estimation of projected seasonal factors does not take into consideration the most recent evolution of the series as in the case of concurrent seasonal factors.

Empirical work at Statistics Canada led to concurrent adjustment of labour force series for the first time, in 1975 using the X-11-ARIMA method. More recent empirical studies have shown the advantages of using concurrent seasonal factors instead of forecast seasonal factors from the viewpoint of revisions (see e.g. Bayer and Wilcox [1981]; Cleveland, Grambsch and Terpenning [1982]; Kenny and Durbin [1982] and McKenzie [1982]).

Further theoretical evidence of the advantage of concurrent seasonal adjustment is given in Dagum [1982c] for the X-11 and X-11-ARIMA methods. However, the benefit of using concurrent seasonal factors is practically null if seasonality is very regular or if the most recent values are strongly contaminated by outliers.

For the series analysed, here, tests have been made to assess which of the two factors is preferable for their current seasonal adjustment.

3.3 The Level of Aggregation Problem

The Consumer Price Index, is a measure of changes in retail prices paid by the target population for consumer goods and services and as such is constructed as a weighted average of price relatives for all single commodities in the "basket." Since the CPI is an aggregate the classical question arises concerning its seasonal adjustment: Is it preferable to seasonally adjust the single commodity price relatives first (if applicable) and aggregate them after using the weights or would seasonal adjustment of the aggregate price index produce better results? The former procedure is referred to in the literature as indirect or derived adjustment while the latter approach is called the direct seasonal adjustment of aggregate series

To further complicate the problem, within the indirect method it is not necessary to seasonally adjust at the single commodity price relative level. It is possible to build the seasonally adjusted All-items CPI from any lower aggregate price indexes, for example, the ones pertaining to major commodity groupings, as long as the appropriate weights are applied.

Of course, the choice of the most appropriate method would not pose a problem if the seasonally adjusted estimates resulting from the different procedures were identical or at least statistically equivalent, but this is not generally the case. As pointed out by Lothian and Morry [1977] the most frequently used seasonal adjustment methods, the method II-X-11 variant and the X-11-ARIMA have inherent non-linearities that will introduce differences between the estimates produced by the direct and the indirect procedures. The non-linearities of these two methods can be eliminated if the following conditions are satisfied:

(1) All the component series as well as the aggregate are seasonally adjusted with the additive decomposition model;

(2) Extreme values are not replaced;

(3) The variable trend-cycle curve (Henderson filter) is kept constant during all the iterations.

he fulfillment of these conditions, however, would generally be attained to the detriment f the seasonal adjustment quality.

Another way of approaching the seasonal adjustment of aggregate series suggested by eweke [1978] makes use of the conditional mathematical expectation of the aggregate ith respect to all the components jointly. This method is a special case of the general eory of prediction in stationary stochastic processes developed by Wold [1938], olmogorov [1939], Wiener [1949] and Whittle [1963].

Although Geweke's approach is theoretically appealing, it suffers from several drawbacks for practical purposes, the most important being the fact that the method requires the knowledge of the joint distributions of the aggregate, the seasonal and the non-seasonal components, conditions that can be satisfied by simulated series only.

In the discussion to follow, the problem of seasonally adjusting an aggregate price index series will be considered in the context of the X-11-ARIMA method.

Conceptually neither the direct nor the indirect procedure is optimal. There are arguments in favour of both approaches.

The direct procedure:

(i) has the advantage of providing historical seasonally adjusted data no matter how often the consumer basket is updated (the seasonal adjustment quality could be affected if there are large shifts in the expenditure weights);

(ii) it is operationally faster and less expensive;

(iii) it offers a possible cancellation effect in the month-to-month movement for seasonal commodities with opposing seasons.

On the other hand, the indirect procedure:

(i) provides the analytical tool of attributing the change in the All-items CPI to changes in the price of different commodities;

(ii) allows for seasonal adjustment at the basic elementary unit level where seasonality is well identifiable, displays a relatively simple pattern – these latter properties were discussed in detail by Dagum [1979].

If the direct and indirect procedures produce estimates that are not too different from each other, then the selection of the one considered "best" should be based on the relative

mportance of the above advantages. However, if the two sets of seasonally adjusted values liffer significantly, preference should be given to the procedure that yields more accurate estimates based on some relevant criteria. For the consumer price indexes the criteria chosen were the size of the revisions and the degree of smoothness of the seasonally adjusted series.

. Evaluation Criteria

Smoothness in a seasonally adjusted series is generally a desirable property to policy nakers. It is expected that removal of the seasonal component eliminates a source of recurring variability in the series, and thus, that the seasonally adjusted series be less rough han the original. Smoothness, however, should not be the sole criterion by which to judge n adjustment procedure for it is always possible to have a highly smoothed seasonally djusted series by the use of seasonal factors that fluctuate less regularly. In fact, fitting easonal factors sufficiently closely to the data one can achieve nearly any degree of noothness in the adjusted series. To avoid this, a second criterion must be considered, e size of the revisions of the seasonal factors. The minimization of revisions in seasonal-adjusted data has received much current attention. Revisions arise under the current ethod of seasonal adjustment, the X-11-ARIMA, because of differences in the asymmetric filters applied to the most recent years of data and the innovations that enter into e series with new observations.

It is important for policy making that the revisions of current seasonally adjusted series as small as possible. Because a current seasonally adjusted value can be obtained by plying a concurrent seasonal factor or a projected seasonal factor, the revision measures e defined below for both cases.

1 Measures of Revision

As already stated in Section 3.1, revisions of the seasonally adjusted series by X-11 and -11-ARIMA stop when the central symmetric filters are applied. In the context of these o methods, this means that the first and last three and a half years will be revised because good approximation of their filters requires 85 points of monthly data. The total revision of a concurrent seasonal factor $\hat{S}_t^c$ is defined by

$$R(\hat{S}_t^c) = \hat{S}_t^c - \hat{S}_t^f \tag{4.1.1}$$

Where $\hat{S}_t{}^F$ is the "final" seasonal factor, in the sense that it will no longer change when the series is augmented with new data. Similarly for the concurrent seasonally adjusted series denoted by $\hat{CI}_t{}^c$, the total revision measure is[1]

$$R(C\hat{I}_t^c) = C\hat{I}_t^c - C\hat{I}_t^f \tag{4.1.2}$$

Substituting the projected seasonal factor $\hat{S}_t{}^p$ for $\hat{S}_t{}^c$ in equation (4.1.1) we obtain a revision measure for the projected seasonal factor and similarly for the current seasonally adjusted series in equation (4.1.2) if a projected seasonal factor is applied.

In the context of X-11-ARIMA at least five years of data are necessary to produce a seasonally adjusted series, we would then need nine and a half years to obtain 12 points t for which the total revisions defined above can be calculated. Because the length of the series analysed in this study is eight years, starting in January 1974 and ending in December 1981, equations (4.1.1) and (4.1.2) cannot be applied (there is no "final" seasonal factor). To overcome the limitations of lack of data points, the procedure followed here consists of comparing each of the concurrent and projected seasonal factors, of the last three years, with the corresponding estimates obtained from the seasonal adjustment of the data up to the end of year 1981, denoted by $\hat{S}_t{}^{81}$. Thus, the modified revision measures for concurrent and projected seasonal factors are,

$$R_1(\hat{S}_t^c) = \hat{S}_t^c - \hat{S}_t^{81} \tag{4.1.3}$$

$$R_1(\hat{S}_t^p) = \hat{S}_t^p - \hat{S}_t^{81} \tag{4.1.4}$$

and similarly, for the current seasonally adjusted series,

$$R_1(C\hat{I}_t^c) = C\hat{I}_t^c - C\hat{I}_t^{81} \tag{4.1.5}$$

$$R_1(C\hat{I}_t^p) = C\hat{I}_t^p - C\hat{I}_t^{81} \tag{4.1.6}$$

The above substitutions are justified because the concurrent and forecasting seasonal filters converge monotonically to the central filter. The actual size of the total revisions, however, will be larger than those obtained with the modified measures.

The revision measures calculated for the series analyzed in this study are the mean absolute revisions of the current seasonally adjusted series obtained with concurrent (c) and projected (p) seasonal factors. That is,

$$|\bar{R}^{c}| = \sum_{t=61}^{96} |C\hat{I}_t^{c} - C\hat{I}_t^{81}| \;/36 \qquad (4.1.7)$$

$$|\bar{R}^{p}| = \sum_{t=61}^{96} |C\hat{I}_t^{p} - C\hat{I}_t^{81}| \;/36 \qquad (4.1.8)$$

The procedure giving the smallest mean absolute revision is to be preferred.

4.2 Measures of Smoothness

In the context of consumer price index series, the level is not as important as the month-to-month percentage changes; otherwise referred to as growth rate. Thus the smoothness of the seasonally adjusted data will be measured in terms of its growth rate r_t.

The measure used is the following:

$$S = \sqrt{\sum_{t=62}^{96} (r_t - \bar{r})^2 \;/35} \qquad (4.2.1)$$

$$\text{where} \quad r_t = \frac{C\hat{I}_t - C\hat{I}_{t-1}}{C\hat{I}_{t-1}} \times 100 \qquad (4.2.2)$$

$$\text{and} \quad \bar{r} = \sum_{t=62}^{96} r_t \;/35 \qquad (4.2.3)$$

S is a measure of dispersion about the average growth rate during the last three years. The seasonally adjusted estimates in equation (4.2.2) can refer to concurrent, projected or 1981 December estimates. When two seasonal adjustment procedures produce estimates with similar S values, the measure S is complemented by the measure 'D' designed to detect persistance of growth rate level and defined by,

$$D = \sum_{t=63}^{96} | r_t - r_{t-1} | \, /34 \tag{4.2.4}$$

Preference will be given to the method with lower D values.

4.3 Seasonal Factors and Growth Rate Revisions

It is possible to measure the effect that revisions in the seasonal factor estimates have on the monthly growth rates of price indexes (see Maravall, 1981). Assuming that one-month lag is a relatively small unit, equation (4.2.2) may be written as

$$r_t = \frac{d\, C\hat{I}_t}{C\hat{I}_t} = d \log C\hat{I}_t \simeq \log C\hat{I}_t - \log C\hat{I}_{t-1} \tag{4.3.1}$$

where d symbolizes the differential operator.

The seasonally adjusted series $C\hat{I}_t$ is equal to the original series X_t divided by the estimated seasonal factor $\hat{S}_t$; that is

$$C\hat{I}_t = X_t \,/\, \hat{S}_t \tag{4.3.2}$$

Therefore, substituting equation (4.3.2) in equation (4.3.1) it becomes,

$$r_t \simeq \Delta\log X_t - \Delta\log \hat{S}_t \tag{4.3.3}$$

where Δ is the difference operator.

Denoting the revision error of S_t as ϵ_t, such that

$$\hat{S}_t = S_t + \epsilon_t \tag{4.3.4}$$

hen,

$$\log \hat{S}_t = \log S_t \left(1 + \frac{\epsilon_t}{S_t}\right)$$

$$= \log S_t + \log \left(1 + \frac{\epsilon_t}{S_t}\right) \tag{4.3.5}$$

ince ϵ_t/S_t is a small number; using Taylor expansion an approximation of equation (4.3.5) ay be written as,

$$\log \hat{S}_t \simeq \log S_t + \frac{\epsilon_t}{S_t} \tag{4.3.6}$$

ubstituting equation (4.3.6) into equation (4.3.3),

$$r_t \simeq \Delta\log X_t - \Delta\log S_t - v_t \tag{4.3.7}$$

here

$$v_t = \frac{\epsilon_t}{S_t} - \frac{\epsilon_{t-1}}{S_{t-1}} \tag{4.3.8}$$

us, the variance of r_t due to the seasonal factor revision ϵ_t is given by

$$\mathrm{var}_S(r_t) = \mathrm{var}_S\left(\frac{\epsilon_t}{S_t} - \frac{\epsilon_{t-1}}{S_{t-1}}\right) \tag{4.3.9}$$

The variance depends on S. If the seasonal factors are near one and assuming no error the original series X_t, then

$$\mathrm{var}\ r_t \simeq 2\ \mathrm{var}\ \epsilon_t\ (1-\rho(1)) \tag{4.3.10}$$

Equation (4.3.10) states that the effect of the seasonal factor error on the month-to-month rate of change can be calculated from the variance of the error series and its firs order autocorrelation.

5. Selecting the Appropriate Seasonal Adjustment Procedure for the Canadia Consumer Price Index Series

There were three major areas to investigate concerning the seasonal adjustment of th Canadian Consumer Price Index series.

(1) Does ARIMA extrapolation improve the quality of the seasonally adjusted estimates

(2) Is the use of concurrent seasonal factors preferable to using forecast factors to pro duce current seasonally adjusted estimates?

(3) Is indirect adjustment better than the direct procedure and if yes what level of disag gregation is optimal for the seasonal adjustment of the All-items Consumer Pric Index series?

Since testing the first two problems would proceed very differently if the direct procedu is adopted instead of the indirect one, it was necessary to start out with examining th third problem first.

5.1. The Level of Aggregation Problem

The data used to resolve the problem of aggregation was supplied by Prices Divisio (The last three years are shown in Appendixes 1 to 3.) It consisted of three sets of seasonal adjusted All-items CPI series extending from January 1974 to December 1981 with 197 as the base year. The first series was the seasonally adjusted All-items Consumer Pri Index obtained as an aggregate of item level price index series, some of which were seasonal adjusted, others unadjusted. Series that contained identifiable seasonality entered the a gregate in an adjusted form while series with no apparent seasonal pattern were left inta before summing up. In this type of indirect procedure, there is always a danger that season movements that go undetected in the low level series due to the presence of high irregul fluctuations get more pronounced through the aggregation procedure (where irregula cancel out) and produce residual seasonality in the seasonally adjusted aggregate. T

eights used in the aggregation were based on the 1974 commodity basket. (The above rocedure coincides with the one used for the calculation of the monthly published rate uring the period of September 1978 to December 1981.)

This official series served as the standard for comparison against two alternative pro-edures; one alternative was seasonally adjusting directly the published All-items CPI series. he second procedure consisted of seasonally adjusting the seven published major com-odity group price index series and summing them up (using the applicable expenditure atterns during the time period under consideration) to obtain the seasonally adjusted All-ems CPI.

To avoid difficulties arising from the non-homogeneous treatment of the three procedures ith regard to linking, testing was carried out on the last three years of data only. This pproach also ensured that the findings pertained to the most recent and consequently most levant movement of the series.

In Section 4, two types of measures were introduced for comparing seasonal adjustment uality. For the study of the aggregation problem, however, the analysis of the three sets estimates is restricted to examining the degree of smoothness only. The calculation of vision measures would have involved thousands of seasonal adjustment runs which was ot considered feasible either operationally or financially.

Table 5.1 presents a summary of the findings. The smoothness measures S and D were lculated based on the last three years of the seasonally adjusted series $\hat{CI}_t^{81}$ obtained th the three different procedures.

According to Table 5.1 there is a marked improvement in smoothness when switching om the officially published indirect adjustment to either of the two alternative procedures. hile seasonally adjusting by the original indirect procedure introduced only 12% reduc-n in the fluctuations of the growth rate, the direct procedure reduced it by 36% (three es as much) as attested by the value .190 versus .296 present in the raw series.

Figure 5.1
Power Spectrum Estimates of Seasonally Adjusted C.P.I. (All Items) Month to Month Percentage Change

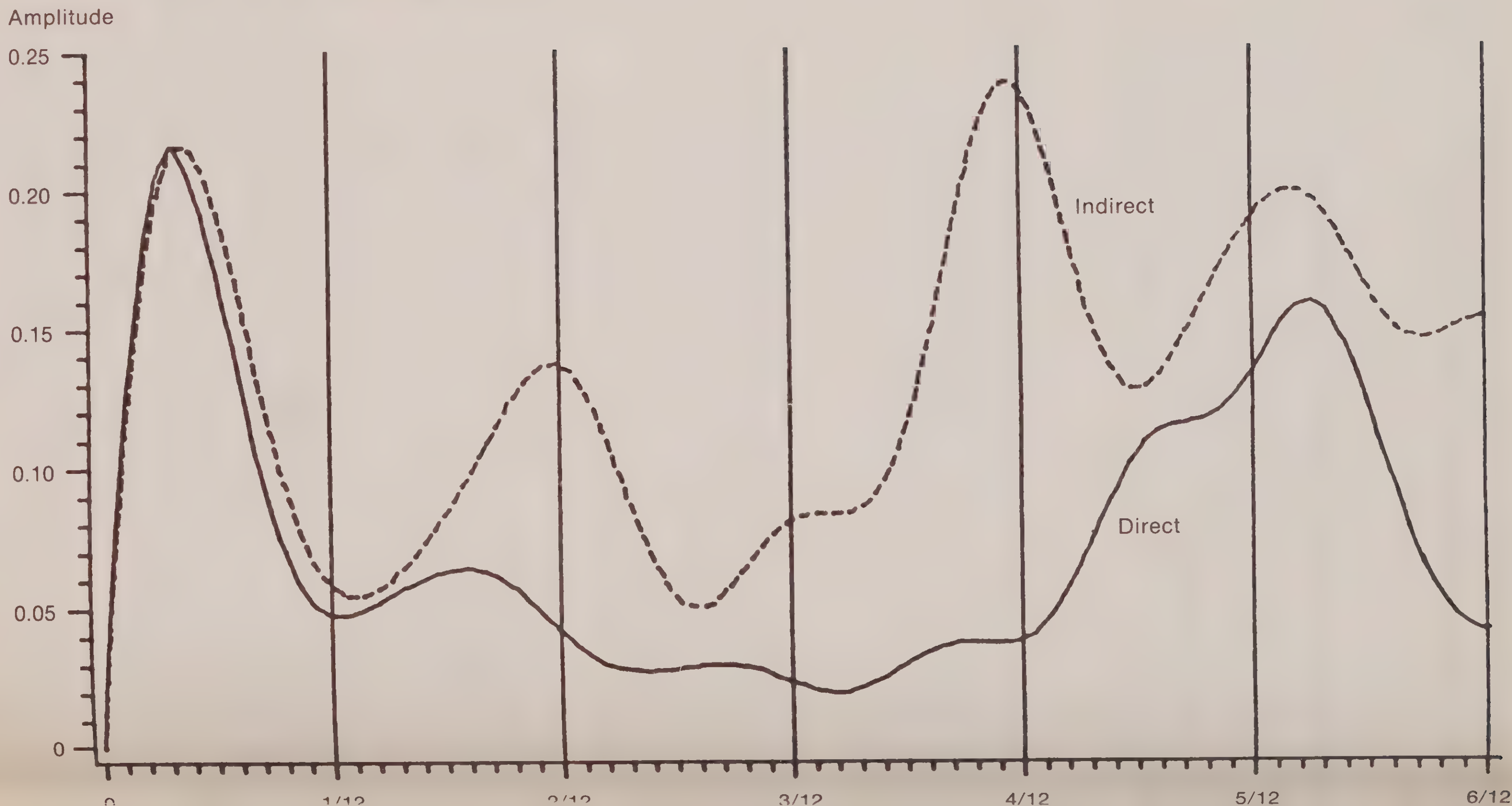


TABLE 5.1 Smoothness Measures of the All-items Consumer Price Index from Three Seasonal Adjustment Procedures

	Unadjusted	Indirect adjustment	Direct adjustment	Indirect adjustment (7 major Sub-groups)
Standard deviation of the growth rate 'S'	.296	.261	.190	.201
Absolute differences in the growth rate 'D'	.330	.324	.212	.226

Comparing the spectrum of the seasonally adjusted series obtained by the direct and indirect methods (Figure 5.1) leads to similar conclusions i.e. the direct method is preferable :o the indirect one. The indirect procedure instead of reducing power at all seasonal fre-quencies (as it is expected of any seasonal adjustment method) even introduces power at he seasonal frequencies 2/12 and 4/12. This is clearly not acceptable.

When the results of this research were made available to Prices Division, a decision was nade to adopt the direct procedure for the seasonal adjustment of the All-items CPI series tarting with the release of the 1982 April index. The direct procedure was favoured over he second alternative since it produced slightly smoother estimates and was also more con-enient operationally. However, if in the future a need arises to attribute movements in he aggregate to changes in major commodity price indexes, then building the CPI from ıe seven seasonally adjusted major subgroups is equally acceptable.

Due to the special importance of two other price index series, Food and All-items-cluding Food, a study was conducted on these two series to indicate whether it is desirable change to direct seasonal adjustment from the previously accepted indirect procedure. the case of the All-items-excluding Food CPI a third procedure was investigated as well; re the seasonally adjusted index was obtained as an aggregate of its six seasonally ad-sted major sub-components.

Tables 5.2 and 5.3 summarize the results of this latter study.

TABLE 5.2 Smoothness Measures of the Consumer Price Index – Food, from two Seasonal Adjustment Procedures

	Unadjusted	Indirect adjustment	Direct adjustment
Standard deviation of the growth rate 'S'	.835	.770	.721
Absolute difference in the growth rate 'D'	.809	.798	.656

TABLE 5.3 Smoothness Measure of the Consumer Price Index – All-items-excluding Food from three Seasonal Adjustment Procedures

	Unadjusted	Indirect adjustment	Direct adjustment	Indirect adjustment (6 major subgroups)
Standard deviation of the growth rate 'S'	.319	.277	.255	.260
Absolute difference in the growth rate 'D'	.393	.329	.281	.273

The findings were similar to those obtained for the All-items Consumer Price Index Direct seasonal adjustment yielded smoother series than the indirect procedure. For the Food CPI, the reduction in fluctuations was 14% with the direct method as opposed to 8% using indirect adjustment, for All-items-excluding Food the corresponding figures were 20% and 13%. For this latter series the gains introduced by the six component indirect procedure were comparable to those originating from the direct approach although slightly smaller.

On the basis of this evidence and following our recommendations, Prices Division changed over to the direct procedure for both series with the release of the 1982 April index.

With the problem of aggregation resolved and the direct approach adopted, our task

of examining the possibility of ARIMA extrapolations and the suitability of concurrent seasonal factors was greatly simplified.

Section 5.2 presents the result of the analysis concerning these two topics.

5.2 Comparison of Four Different Options for Producing Current Seasonally Adjusted Estimates of the Consumer Price Index

Once the preference for the direct approach is established for the seasonal adjustment of the All-items Consumer Price Index series four options must yet be considered in the context of direct adjustment. These options are the use of the X-11-ARIMA method with and without the ARIMA extrapolation and, for each case, the choice of concurrent seasonal factors or forecast seasonal factors when calculating current seasonally adjusted figures.

Regarding the use of ARIMA extrapolation, theoretical studies by one of the authors, [Dagum 1982a and 1982b] showed that the enlargement of the raw series with one year of extrapolated values substantially reduced filter revisions. However, depending on the nature of the series the seasonal adjustment filters are applied to, this reduction in filter distance may or may not result into a reduction in the size of revisions the series will undergo as further data become available. The ARIMA extrapolation produces best results when the series in question conforms to a certain structure concerning seasonal pattern, and irregular variation. Since the Consumer Price Index series displayed rather unusual seasonal structures as shown in Section 2, it is necessary to undertake an empirical study to decide if it is beneficial to extend the series with extrapolated values before seasonal adjustment.

Similarly, theoretical work has been carried out [Dagum 1982c] on the properties of concurrent and forecast X-11-ARIMA seasonal filters with regard to revisions. Although concurrent filters are closer to the final central filters than forecasting filters, thus implying smaller filter revisions, due to the data dependence of the total revisions, especially in the presence of outliers, empirical work is also needed to indicate which factors produce better estimates in the case of the Consumer Price Index series.

The investigation of the appropriate method (X-11-ARIMA with or without extrapolation) and the optimal type of seasonal factors (concurrent or forecast) to use pose a problem of circularity. In order to assess the desirability of ARIMA extrapolations the preferred type of seasonal factors have to be used, on the other hand, to choose between concurrent and forecast factors the question of a suitable method has to be settled first. To avoid difficulties arising from this circularity, the two problems are examined simultaneously. Four sets of seasonally adjusted figures are produced using concurrent seasonal factors obtained with and without ARIMA extrapolation, and forecast seasonal factors originating from an X-11-ARIMA adjustment with and without the use of the ARIMA option.

The optimal seasonal adjustment will be given by the combination of the appropriate method and of seasonal factors that yield the smoothest estimates with the smallest revisions.

The data included in the analysis are again restricted to the last three years, from January 1979 to December 1981, in order to reveal the most recent performance of the four adjustments in question without reducing the sample size too much. Both the smoothness and revision measures of the four seasonally adjusted series are calculated to determine which estimates satisfy best the chosen criteria, as defined in Sections 4.1 and 4.2.

Table 5.4 gives a summary on the results of the analysis.

TABLE 5.4 Smoothness and Revision Measures of the All-items Consumer Price Index from four Seasonal Adjustment Procedures

		X-11-ARIMA without extrapolation		X-11-ARIMA with extrapolation	
Measure	Unadjusted	Concurrent	Forecast	Concurrent	Forecast
S	.296	.272	.271	.237	.269
D	.330	.321	.338	.232	.300
$\overline{\lvert R\rvert}$		.114	.131	.103	.125

From Table 5.4, it is evident that using ARIMA extrapolation and applying concurrent seasonal factors produces the best seasonally adjusted estimates of the four alternatives,

both from the point of view of degree of smoothness and the size of the revisions. The adoption of these combined options for seasonal adjustment will result in a 20% reduction in the growth rate fluctuations as opposed to only 9% in the officially used method (forecast seasonal factors and ARIMA extrapolation). Thus, estimates produced in this nanner will give more accurate month-to-month movements than the official figures.

In terms of revision, the improvement is also significant, the average revision size of 125 present when using the official procedure is lowered to .103 when estimates are obained with the proposed combination. Furthermore, it should be kept in mind, as menioned in Section 4, that these revisions are not the final ones and the values in Table 5.4 nderestimate the size of the total revisions. For the final revisions, the absolute difference etween the two measures would be much larger. Examining data in the year 1979 only, or which the revisions calculated are almost identical to the final ones, the revision measures re .167 and .100 for the two alternative adjustments in question.

Switching to the use of concurrent seasonal factors poses a dilemma concerning the release f month-to-month percentage changes in the Consumer Price Index. Should the month-)-month movement be calculated based on the seasonally adjusted level reported in the evious month or should the revised last month level serve as the basis of percentage change lculations? This question does not arise in the context of forecast seasonal factors since e series are processed through the X-11-ARIMA program only once a year and, from en on, seasonally adjusted data are produced with the help of the projected factors which main unchanged for the duration of the next year.

To solve the above problem we again calculated the smoothness and revision measures roduced earlier. In Table 5.4, the smoothness and revision measures of the growth rate e calculated using **non-revised previous month's values**. Table 5.5 shows these measures tained when the growth rate is based on **first revised last month's estimates**.

According to the entries in Table 5.5 there is a definite gain from the incorporation of st-revised estimates in the calculation of the month-to-month percentage changes both terms of smoothness and in the size of the revision.

TABLE 5.5 Comparison of Preliminary and First-revised Growth Rate Measures of th All-items Consumer Price Index Using Concurrent Factors with ARIMA Extrapolations

Measure	Preliminary Estimates	First-revised Estimates
S	.237	.200
D	232	.176
$\lvert\bar{R}\rvert$	103	.091

The results of the empirical study strongly suggest that:

(1) the seasonally adjusted Consumer Price Index be obtained with the direct procedu or, at most, with indirect adjustment of the seven major subcomponents;

(2) the X-11-ARIMA method be used with the ARIMA extrapolation option;

(3) the current seasonally adjusted estimates be calculated with the aid of concurre seasonal factors instead of seasonal factor forecasts;

(4) the calculation of month-to-month percentage changes be based on the first-revis previous month's estimates.

The implementation of these practices will result in a significantly smoother seasonal adjusted CPI series that will undergo smaller revisions when new observations are inco porated into the series.

. Conclusions

This study has analysed the main characteristics of seasonal variations in consumer price ıdex series and compared several seasonal adjustment procedures for their estimation.

Among the series discussed, the spectrum of the All-items CPI showed seasonal variaons that differ significantly from those observed in most economic indicators. There is ɔ power at the fundamental seasonal band associated with the annual cycle and the domiant peaks correspond to the 4-, 2.4- and 2-month cycles. These seasonal peaks result from e fact that most of the items that enter with large weights are subject to different and me varying frequency of price collection. The time varying property introduces a rapidly ıanging seasonal structure which is reflected in the spectrum by very broad seasonal bands.

The comparison of the seasonal adjustment procedures referred to: (1) the use of the -11-ARIMA seasonal adjustment method with and without the ARIMA extrapolation ɔtion; (2) the use of concurrent versus year-ahead seasonal factors for current seasonal ljustment; and (3) the level of aggregation at which the seasonal adjustment must be done. ıe selection of the "best" procedure was made based on two highly desirable properties, e minimization of the revisions and the maximization of the degree of smoothness of e seasonally adjusted series.

The empirical results of Section 5 strongly indicated that the preferred seasonal adjustent procedure for the All-items Consumer Price Index series consists of:

1) direct seasonal adjustment of the aggregated series (being the second best the indirect seasonal adjustment with the seven major subcomponents);

2) the use of the ARIMA extrapolation of the X-11-ARIMA method.

3) the use of concurrent seasonal factors for current seasonal adjustment; and

4) the calculation of month-to-month changes based on the first-revised previous month's estimates.

In selecting the direct seasonal adjustment of the published All-items CPI, however, w should always take into consideration the effect that the updating of fixed baskets ma have on the seasonal structure.

During the period analysed, no significant abrupt changes were observed, possibly becau there were only two expenditure patterns in effect during the period studied. The frequer cy of weight revisions that would minimize abrupt changes in the seasonal component an the use of variable versus constant weights for seasonal items are very important topi that need further research.

Another relevant issue not covered, here, concerns the data reporting practices. Variabilit in growth rates in economic series depends in part on the frequency of data reporting (da month, quarter and so on) and it part on the length of the interval over which th change is measured.

nerally, the variability of the rates diminishes as the span over which the change i asured increases to, say, three months, six months or 12 months.

Methods of measuring growth rates can be devised to take advantage of the reductio in variance both as the reporting frequency decreases and as the interval between unit is lengthened.

On the other hand, rates of growth over shorter spans have the advantage of registerin turning points promptly if the series are smooth. But if the variability of the series is high short-span rates contain many false signals and time is required to separate the true fron the false. Another alternative is to consider ''filtered'' seasonally adjusted series, wher the filter has removed most of the irregular variations leaving only the trend-cycle fluctua tions. Research on these important issues is definitely needed and should be strongl encouraged.

Footnote

[1] We make the assumption that the decomposition model for the time series X is multiplicative, that is, X = CSI where C denotes the trend cycle, S the seasonal factor and I the irregular fluctuation.

References

Akaike, H. and Ishiguro, M. [1980]: ''BAYSEA, a Bayesian Seasonal Adjustmen Program'', Computer Science Monograph No. 13, The Institute of Statistica Mathematics, Tokyo, Japan.

Bayer, A. and Wilcox, D. [1981]: ''An Evaluation of Concurrent Seasonal Adjustment'' Technical Report, Special Studies Section, Washington, D.C.: Board of Governors o the Federal Reserve System.

Bilongo, R. and Carbone, R. [1981]: ''Seasonal Adjustment and Adaptive Forecasting' Research Paper, Faculté des Sciences de l'Administration, Université Laval, Québec Canada.

Box, G.E.P. and Jenkins, G.M. [1970]: *Time Series Analysis: Forecasting and Control* San Francisco: Holden Day.

Burman, J.P. [1980]: ''Seasonal Adjustment by Signal Extraction'', *Journal of the Roya Statistical Society*, Series A, 143, Part 3, pp.321-27.

Cleveland, W.S., Grambsch P. and Terpenning, I. [1982]: ''A Comparison of Concurren and Forecast Seasonal Adjustment'', in *Proceedings of the Business and Economic Section*, Annual Meetings of the American Statistical Association (forthcoming).

Dagum, E.B. [1975]: ''Seasonal Factor Forecasts from ARIMA Models'', *Proceedings o the International Statistical Institute*, 40th Session, Vol. 3, pp.206-19.

__________ [1978]: *Comparison and Assessment of Seasonal Adjustment Methods fo Labour Force Series*, Washington, D.C.: U.S. Government Printing Office.

__________ [1979]: *The Seasonal Adjustment of Economic Time Series Aggregates. A Case Study: The Unemployment Rate*, Washington, D.C.: U.S. Government Printin Office.

———— [1980]: *The X-11-ARIMA Seasonal Adjustment Method*, Ottawa: Statistics Canada, Catalogue No. 12-564E.

———— [1982a]: "Revision of Time Varying Seasonal Filters", *Journal of Forecasting*, Vol. 1, pp.173-87.

———— [1982b]: "The Effects of Asymmetric Filters on Seasonal Factor Revisions", *Journal of the American Statistical Association*, Vol. 77, No. 380, pp.732-738.

———— [1982c]: "Revisions of Seasonally Adjusted Data Due to Filter Changes", in *Proceedings of the Business and Economics Section*, Annual Meetings of the American Statistical Association, pp.39-45.

)uvall, R.M. [1966]: "Time Series Analysis by Modified Least-Squares Techniques", *Journal of the American Statistical Association*, Vol. 61, pp.152-65.

'arley, D. and Zeller, S.H. [1978]: "Comments", in *Seasonal Analysis of Economic Time Series*, (Arnold Zellner, Editor), Washington, D.C.: U.S. Government Printing Office, pp.464-67.

iersovitz, M. and MacKinnon, J.G. [1978]: "Seasonality in Regression: An Application of Smoothness Priors", *Journal of the American Statistical Association*, Vol. 73, pp.264-73.

eweke, J. [1978]: "Temporal and Sectoral Aggregations of Seasonally Adjusted Time Series", in *Seasonal Analysis of Economic Time Series,* (Arnold Zellner, Editor), Washington, D.C.: U.S. Government Printing Office, pp.411-427.

ranger, C.W.J. [1978]: "Seasonality: Causation, Interpretation and Implications", in *Seasonal Analysis of Economic Time Series,* (Arnold Zellner, Editor) Washington, D.C.: U.S. Government Printing Office, pp.33-46.

Iannan, E.J. [1960]: "The Estimation of Seasonal Variations", *The Australian Journal of Statistics*, Vol. 2, pp.1-15.

Havenner, A. and Swamy, P.A.V.B. [1978]: "A Random Coefficient Approach to Seasona Adjustment of Economic Time Series", Special Studies Paper 124, Research and Statistic Division, Washington, D.C.: Board of Governors of the Federal Reserve System.

Hatanaka, M. and Suzuki, M. [1967]: "A Theory of the Pseudo-spectrum and it Applications to Non-Stationary Dynamic Econometric Models", in *Essays i Mathematical Economics in Honor of Oskar Morgenstern* (Martin Shubik, Editor) New Jersey: Princeton University Press.

Henshaw, R.C. [1966]: "Application of the General Linear Model to Seasonal Adjus ment of Economic Time Series", *Econometrica*, Vol. 34, pp.381-95.

Hillmer, S.C. and Tiao, G.C. [1982]: "An ARIMA-Model Based Approach to Seasona Adjustment", *Journal of the American Statistical Association*, Vol. 77, No. 377, pp.63-70.

Jorgenson, D.W. [1964]: "Minimum Variance Linear Unbiased Seasonal Adjustment of Economic Time Series", *Journal of the American Statistical Association*, Vol. 59, pp.681-723.

Kenny, P. and Durbin, J. [1982]: "Local Trend Estimation and Seasonal Adjustment of Economic Time Series", *Journal of the Royal Statistical Society*, Series A, 145, Part 1, pp.1-41.

Kolmogorov, A.N. [1939]: "Sur l'interpolation des suites stationnaires", Comptes Rendues de l'Academie de Sciences, Paris, Vol. 208, p.2043.

Kuiper, J. [1981]: "The Treatment of Extreme Values in the X-11-ARIMA Program", *Time Series Analysis and Forecasting*, (Anderson, O. and Perryman, M.R., Editors) Amsterdam: North-Holland Publishing Co., pp.257-66.

Ladd, G.W. [1964]: "Regression Analysis of Seasonal Data", *Journal of the American Statistical Association*, Vol. 59, pp.402-21.

Lothian, J. and Morry, M. [1977]: "The Problem of Aggregation: Direct or Indirect Seasonal Adjustment", Research Paper, Time Series Research and Analysis Division, Ottawa: Statistics Canada.

Lovell, M.C. [1963]: "Seasonal Adjustment of Economic Time Series and Multiple Regression Analysis", *Journal of the American Statistical Association*, Vol. 58, pp.993-1010.

Maravall, A. [1981]: "Errores de Medicion del Crecimiento a Corto Plazo de Series Monetarias Desestacionalizadas", Servicio de estudios Economicos No. 21, Madrid: Banco de Espa-74-a.

McKenzie, S. [1982]: "An Evaluation of Concurrent Adjustment on Census Bureau Time Series", in *Proceedings of the Business and Economics Section*, Annual Meetings of the American Statistical Association (forthcoming).

Nerlove, M. [1964]: "Spectral Analysis of Seasonal Adjustment Procedures", *Econometrica*, Vol. 32, pp.241-286.

———— Grether, D.M. and Carvalho, J.L. [1979]: *Analysis of Economic Time Series*, New York: Academic Press.

Pierce, D. [1978]: "Seasonal Adjustment When Both Deterministic and Stochastic Seasonality are Present", in *Seasonal Analysis of Economic Time Series*, (Arnold Zellner, Editor), Washington, D.C.: U.S. Government Printing Office, pp.242-269.

———— [1980]: "Data Revision with Moving Average Seasonal Adjustment Procedures", *Journal of Econometrics*, Vol. 14, No. 1, pp.95-114.

Rosenblatt, H.M. [1963]: "Spectral Analysis and Parametric Methods for Seasonal Adjustment of Economic Time Series", *Proceedings of the Business and Economics Section of the Annual Meeting of the American Statistical Association*, p.94.

Shiskin, J., Young, A.H. and Musgrave, J.C. [1967]: "The X-11 Variant of Census Metho
II Seasonal Adjustment Program", Technical Paper No. 15, Washington, D.C.: U.S
Bureau of Census.

Statistics Canada [1982], *The Consumer Price Index Reference Paper*, Concepts an
Procedures, Catalogue No. 62-553E.

Stephenson, J.A. and Farr, H.T. [1972]: "Seasonal Adjustment of Economic Data b
Application of the General Linear Statistical Model", *Journal of the American Statistic
Association*, Vol. 67, pp.37-45.

Wallis, K.F. [1974]: "Seasonal Adjustment and Relations between Variables", *Journa
of the American Statistical Association*, Vol. 69, pp.18-31.

Whittle, P. [1963]: *Prediction and Regulation by Linear Least-Squares Methods*, London
English University Press.

Wiener, N. [1949]: *Extrapolation, Interpolation and Smoothing of Stationary Time Series*
New York: John Wiley.

Wold, H. [1938]: *A Study in the Analysis of Stationary Time Series*, Stockholm: Almquis
and Wiksell (2nd edition, 1954).

Young, A.H. [1968]: "Linear Approximations to the Census and BLS Seasonal Adjust-
ment Methods", *Journal of the American Statistical Association*, Vol. 63, pp.445-57

APPENDIX 1

All-items Consumer Price Index Unadjusted Series

	Jan.	Feb.	Mar.	Apr.	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.
1979	146.2	147.5	149.3	150.3	151.8	152.5	153.7	154.2	155.6	156.7	158.2	159.2
1980	160.1	161.4	163.2	164.2	166.1	167.9	169.2	170.8	172.3	173.8	176.0	177.0
1981	179.3	181.1	183.5	184.9	186.6	189.4	191.1	192.5	193.9	195.8	197.5	198.4

All-items Consumer Price Index Seasonally Adjusted Series Using the Direct Procedure

	Jan.	Feb.	Mar.	Apr.	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.
1979	146.7	147.9	149.2	150.5	151.6	152.1	153.3	154.0	155.6	156.8	158.1	159.6
1980	160.5	161.8	163.0	164.4	165.9	167.5	168.7	170.6	172.4	173.9	175.9	177.5
1981	179.8	181.5	183.3	185.1	186.4	188.9	190.6	192.3	194.0	195.9	197.3	199.0

All-items Consumer Price Index Seasonally Adjusted Series Using the Indirect Procedure

	Jan.	Feb.	Mar.	Apr.	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.
1979	142.6	143.9	145.7	146.7	147.9	148.5	149.4	150.1	151.9	152.9	154.2	155.4
1980	156.2	151.5	159.2	160.3	161.8	163.5	164.5	166.2	166.2	169.7	171.6	172.9
1981	175.0	176.7	179.2	180.6	181.8	184.5	185.7	187.4	189.3	191.1	192.4	193.6

All-items Consumer Price Index Seasonally Adjusted Series Using Seven Component Indirect Procedures

	Jan.	Feb.	Mar.	Apr.	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.
1979	146.5	147.7	149.1	150.5	151.4	152.0	153.3	154.0	155.6	156.5	158.0	159.3
1980	160.4	161.6	163.0	164.2	165.8	167.5	168.7	170.6	172.2	173.8	175.6	177.3
1981	179.8	181.4	183.2	184.9	186.4	188.9	190.6	192.3	193.8	195.7	197.1	198.6

Consumer Price Index – Food Unadjusted Series

	Jan.	Feb.	Mar.	Apr.	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.
1979	153.6	157.3	161.3	162.8	163.7	164.5	167.4	166.2	166.5	167.6	168.1	170.4
1980	170.8	173.2	174.8	175.4	177.0	181.0	182.6	185.1	188.1	188.9	191.0	193.2
1981	194.2	197.5	198.8	200.8	199.9	203.5	206.2	206.8	206.3	206.1	205.7	204.0

Consumer Price Index – Food Seasonally Adjusted Series Using the Direct Procedure

	Jan.	Feb.	Mar.	Apr.	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.
1979	155.0	157.9	161.8	163.3	163.2	163.2	165.0	164.7	166.2	168.2	169.4	171.5
1980	172.5	173.7	175.1	175.7	176.8	179.7	180.1	183.6	187.6	189.4	192.2	194.5
1981	196.2	198.0	199.1	201.1	200.0	202.2	203.6	205.2	205.7	206.5	206.7	205.4

Consumer Price Index – Food Seasonally Adjusted Series Using the Indirect Procedure

	Jan.	Feb.	Mar.	Apr.	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.
1979	151.9	155.1	158.7	160.3	160.7	160.7	162.1	161.5	163.3	165.0	165.9	168.4
1980	169.0	170.8	172.1	172.6	173.8	176.8	176.8	179.7	184.8	186.1	188.7	191.1
1981	192.1	194.6	195.8	197.8	196.6	198.8	199.6	201.2	202.3	202.7	203.0	201.6

APPENDIX 3

Consumer Price Index – All-items-excluding Food Unadjusted Series

	Jan.	Feb.	Mar.	Apr.	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.
1979	143.4	144.2	145.3	146.1	147.7	148.4	149.2	150.3	151.9	153.0	158.8	155.3
1980	156.3	157.5	159.3	160.4	162.2	163.5	164.8	166.1	167.1	168.9	171.1	171.8
1981	174.4	175.8	178.5	179.7	182.0	184.7	186.1	187.7	189.6	192.1	194.4	195.9

Consumer Price Index – All-items-excluding Food Seasonally Adjusted Series Using the Direct Procedure

	Jan.	Feb.	Mar.	Apr.	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.
1979	143.5	144.4	145.1	146.2	147.6	148.4	149.4	150.5	152.1	152.8	154.2	155.3
1980	156.4	157.7	159.1	160.6	162.1	163.5	164.9	166.4	167.3	168.7	170.4	171.8
1981	174.5	176.1	178.3	179.9	181.8	184.7	186.2	188.0	189.9	192.0	193.6	195.0

Consumer Price Index – All-items-excluding Food Seasonally Adjusted Series Using the Indirect Procedure

	Jan.	Feb.	Mar.	Apr.	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.
1979	140.1	140.8	142.1	143.0	144.3	145.1	145.9	147.0	148.8	149.7	151.0	151.8
1980	152.7	153.8	155.7	156.9	158.5	159.9	161.1	162.5	163.7	165.0	166.9	168.0
1981	170.3	171.8	174.6	175.8	177.8	180.6	181.9	183.6	185.8	187.9	189.5	191.4

Consumer Price Index – All-items-excluding Food Seasonally Adjusted Series Using the Six Component Indirect Procedure

	Jan.	Feb.	Mar.	Apr.	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.
1979	143.6	144.4	145.1	146.4	147.7	148.5	149.6	150.6	152.1	152.8	154.3	155.3
1980	156.4	157.9	159.0	160.5	162.2	163.6	164.8	166.3	167.4	168.7	170.4	171.9
1981	174.6	176.2	178.3	179.8	181.9	184.6	186.4	188.1	189.8	192.2	193.7	196.1

All-item Consumer Price Index Seasonally Adjusted Series Using ARIMA Extrapolation and Concurrent Seasonal Factors

	Jan.	Feb.	Mar.	Apr.	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.
1979	146.6	147.9	149.2	150.7	151.6	152.2	153.0	153.9	155.7	156.6	158.1	159.4
1980	160.7	161.9	163.0	164.3	165.6	167.5	168.5	170.4	172.2	173.8	175.8	177.1
1981	179.7	181.4	183.3	185.0	186.4	188.9	190.6	192.3	194.0	196.0	197.5	199.0

All-items Consumer Price Index Seasonally Adjusted Series Using ARIMA Extrapolation and Forecast Seasonal Factors

	Jan.	Feb.	Mar.	Apr.	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.
1979	146.7	148.0	149.3	150.8	151.6	152.2	153.0	153.9	155.8	156.5	158.0	159.3
1980	160.6	161.8	163.0	164.5	165.7	167.6	168.6	170.6	172.4	174.0	176.0	177.3
1981	179.9	181.6	183.4	185.3	186.3	189.0	190.7	192.3	193.9	195.9	197.1	198.5

All-items Consumer Price Index Seasonally Adjusted Series Using No Extrapolation and Concurrent Seasonal Factors

	Jan.	Feb.	Mar.	Apr.	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.
1979	146.5	147.5	149.4	150.9	151.8	152.4	153.2	153.9	155.6	156.5	158.2	159.5
1980	160.7	161.8	163.0	164.4	165.7	167.6	168.5	170.6	172.4	173.9	176.0	177.2
1981	179.9	181.6	183.4	185.0	186.2	188.9	190.6	192.3	193.9	196.1	197.4	198.9

All-items Consumer Price Index Seasonally Adjusted Series Using No Extrapolation and Forecast Seasonal Factors

	Jan.	Feb.	Mar.	Apr.	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.
1979	146.5	147.8	149.5	150.9	151.8	152.4	153.3	153.9	155.6	156.4	157.9	159.2
1980	160.6	161.8	163.1	164.5	165.7	167.6	168.6	170.6	172.3	173.8	176.0	177.3
1981	179.9	181.6	183.4	185.2	186.2	188.9	190.5	192.3	193.9	196.0	197.4	198.6

All-Items Consumer Price Index Seasonally Adjusted Series Using ARIMA Extrapolation and Concurrent Seasonal Factors First-revised Figures

	Jan.	Feb.	Mar.	Apr.	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.
1979	146.7	147.9	149.3	150.5	151.4	152.1	153.0	154.0	155.6	156.8	158.2	159.5
1980	160.6	161.8	163.0	164.2	165.7	167.3	168.7	170.5	172.2	173.9	175.7	177.4
1981	179.6	181.4	183.2	184.9	186.2	188.8	190.6	192.2	194.0	195.9	197.3	

L'ESTIMATION DES VARIATIONS SAISONNIÈRES DANS LES INDICES DE PRIX À LA CONSOMMATION

To provide you with a version in the official language of your choice, the French text is preceeded by the English text (p.919) in this publication.

Estela Bee Dagum et
Marietta Morry
Division des séries chronologiques, recherche et analyse
Statistique Canada

RÉSUMÉ

Cette étude analyse les caractéristiques des variations saisonnières dans l'Indice des prix à la consommation de tous les articles et ses deux principales composantes, l'alimentation et l'ensemble des articles hors alimentation. Elle évalue laquelle des deux méthodes habituelles de correction des variations saisonnières, la méthode directe ou la méthode indirecte, doit être préférée du point de vue de la régularité du taux moyen d'augmentation. D'autres problèmes connexes étudiés dans ce document sont l'application de la méthode de désaisonnalisation X-11-ARMMI avec ou sans l'option d'extrapolation ARMMI et les facteurs concourants ou prévus. La méthode optimale est choisie en fonction du degré de régularité et de l'ampleur des révisions des séries désaisonnalisées.

Introduction

L'un des usages importants des indices de prix à la consommation désaisonnalisés est l'obtention des taux de croissance de ces statistiques. Le gouvernement peut vouloir fixer des objectifs de croissance afin de guider sa politique, objectifs auxquels sont ensuite comparés les taux effectifs d'augmentation, après élimination des variations saisonnières. La correspondance entre les taux effectifs et les taux visés dépend en partie de la qualité de la correction des variations saisonnières. Des erreurs dans cette correction entraînent des inexactitudes dans les taux d'augmentation désaisonnalisés, ce qui conduit à des incertitudes sur les degrés de réalisation des objectifs. Aussi les révisions apportées aux statistiques

ésaisonnalisées courantes ont-elles toujours préoccupé les responsables de la politique, ı particulier lorsque les révisions sont importantes ou qu'elles modifient le sens de l'évolu-ɔn conjoncturelle.

Un autre caractère souhaitable pour les responsables de la politique est la régularité du ·ofil des données désaisonnalisées. On s'attend à ce que l'élimination de la composante isonnière supprime une source de variabilité récurrente dans les statistiques et donc à que les chiffres désaisonnalisés soient moins irréguliers que les données de départ. Cette ude fait appel à ces deux critères, le degré de régularité et l'ampleur des révisions, pour :terminer la méthode optimale de désaisonnalisation pour l'Indice des prix à la consom-ation de tous les articles et ses deux principales composantes.

Le deuxième chapitre présente les principales caractéristiques des variations saisonnières ns la série de l'Indice des prix à la consommation. Le chapitre 3 expose les diverses éthodes de correction des variations saisonnières, les modèles de régression comparés x modèles à moyenne mobile, les facteurs saisonniers concourants ou prévus et la désaison-lisation directe ou indirecte des séries agrégatives. Le chapitre 4 fournit un ensemble mesures à l'égard des deux critères d'évaluation choisis dans cette étude, le degré de gularité et l'ampleur des révisions apportées aux statistiques désaisonnalisées. Le chapitre ınalyse les résultats empiriques obtenus pour l'IPC de tous les articles et ses deux prin-ıales composantes, l'alimentation et l'ensemble des articles hors alimentation. Quant au apitre 6, il expose les principales conclusions de l'étude.

Variations saisonnières dans les indices de prix à la consommation

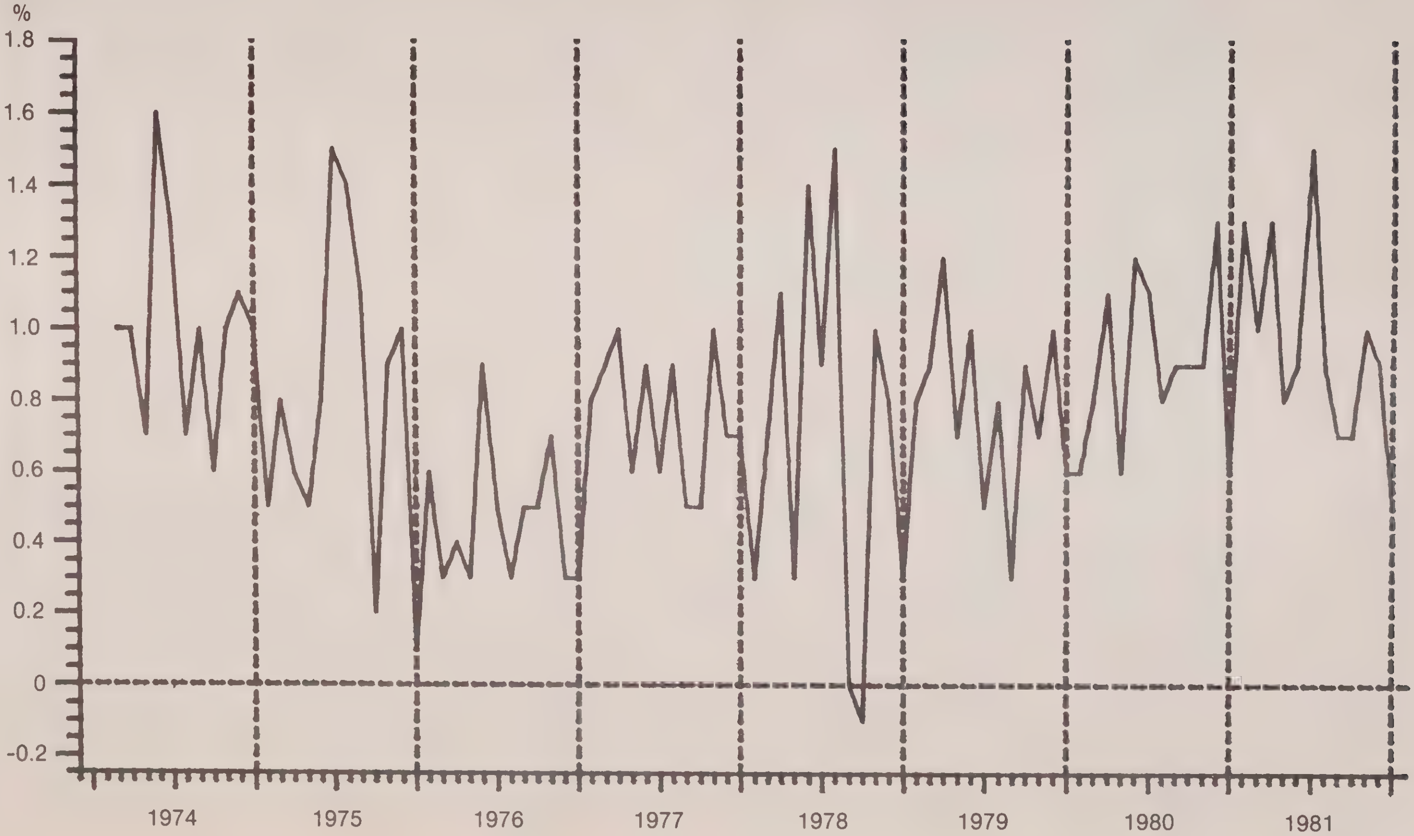
)ans l'analyse des indices mensuels de prix à la consommation, il est extrêmement im-rtant d'estimer l'effet des variations saisonnières. La majeure partie du changement con-é dans des comparaisons d'un mois à l'autre peut être due à un caractère saisonnier ltôt qu'à d'autres mouvements sous-jacents comme une tendance ou un cycle.

.es variations saisonnières représentent l'effet synthétique d'événements climatiques et ıtutionnels qui reviennent avec plus ou moins de régularité chaque année. Dans le cas ticulier des indices de prix à la consommation, la saisonnalité résulte principalement deux causes, le climat (qui influe sur la demande et l'offre de produits) et la fréquence

du relevé des prix, qui dépend de la nature du produit. Mis à part le climat, d'autres facteurs peuvent modifier la saisonnalité des achats de certains produits. Du côté de l'offre, l'évolution des techniques et l'expansion du commerce extérieur peuvent réduire sensiblement l'effet du climat sur l'accessibilité des biens et services. Du côté de la demande, les nouveaux goûts des consommateurs et le revenu de la population peuvent modifier le profil saisonnier de la consommation. Qu'elle soit imputable à une variation de l'offre ou de la demande, une baisse saisonnière des achats d'un produit saisonnier (par exemple les fraises, le maïs, les pardessus d'hiver ou les bicyclettes) est parfois si marquée qu'on ne peut disposer d'un échantillon représentatif des prix de ce produit pendant la morte-saison. Si les coefficients sont fixés pour tous les mois de l'année, les variations de prix doivent être imputées aux produits saisonniers pendant la morte-saison. La méthode employée à Statistique Canada consiste à reporter le dernier prix disponible en saison pendant les mois de la morte-saison. Ainsi, la variation mensuelle relative au premier mois de la saison représente en fait la variation de prix entre ce mois-là et le dernier mois de la saison précédente. Cette méthode présente l'avantage de recourir exclusivement aux prix observés, tout en préservant le principe du panier fixe. Elle a cependant l'inconvénient, en cas de tendance à la hausse des prix, de sous-estimer la variation mensuelle globale des prix en morte-saison et de la surestimer pour le premier mois de la saison. L'autre grand facteur qui introduit un caractère saisonnier, mais pour des produits qui ne sont pas vraiment saisonniers, est la fréquence du relevé des prix. Cette fréquence dépend de la nature du produit considéré. Les biens et services dont les prix changent souvent nécessitent un relevé plus fréquent des prix. Depuis mai 1982, les prix des produits alimentaires sont recueillis deux fois par mois. Pour la plupart des autres produits, on recueille les prix chaque mois, notamment pour les fournitures ménagères, les articles d'ameublement, les vêtements, l'essence, les produits pharmaceutiques, les produits d'hygiène personnelle, les produits du tabac, les boissons alcooliques, les loyers, les intérêts hypothécaires et les habitations neuves. Cependant, le prix des autres produits est recueilli avec une périodicité dépassant le mois; la méthode actuelle consiste à répéter le dernier prix disponible jusqu'à ce qu'on en recueille un nouveau. Le prix des appareils électro-ménagers est recueilli six fois par an. Celui des automobiles, des services d'habillement et des services de soins personnels est recueilli chaque trimestre; celui des journaux, deux fois par an; les droits d'immatriculation des automobiles et les impôts fonciers, annuellement. De plus, les prix de ces produits sont enregistrés à d'autres occasions lorsqu'une modification appréciable de prix se manifeste entre les dates normales d'établissement de leur prix (voir le *Document de référence de l'Indice des prix à la consommation,*

Figure 2.1

IPC (tous articles) non corrigé — Pourcentages de variation mensuelle



concepts et procédés, 1982). Cette fréquence variable de la consignation des prix introdu des variations saisonnières qui suivent un profil nettement différent de celui affiché pa les produits dont le prix est recueilli chaque mois.

Un simple examen de la série donne généralement une bonne idée du genre de saison nalité qu'elle présente. Cependant, dans le cas des variations mensuelles (en pourcentage de l'IPC de l'ensemble des articles, la présence d'un caractère saisonnier est difficile discerner, de sorte que nous recourrons à la technique de l'analyse spectrale.

Le spectre d'une série permet de tirer des renseignements sur les principales caractéristi ques des variations saisonnières influant sur la série en question. Bien que ce spectre soi strictement fondé sur des hypothèses de stationnarité et que la plupart des chronique économiques ne soient pas stationnaires, ce problème peut être éliminé dans le cas d l'analyse de données réelles grâce au concept du pseudo-spectre (Hatanaka et Suzuki, 1967) Le pseudo-spectre est essentiellement le spectre estimé par l'ordinateur comme si les don nées étaient stationnaires. Il peut être envisagé comme la moyenne d'un spectre évoluan dans le temps.

Si le caractère saisonnier est parfaitement périodique, au sens où $S_t = S_{t-12}$, il peu être représenté dans le domaine de fréquences par le modèle simple suivant:

$$S_t = \sum_{j=1}^{6} (\alpha_j \cos 2\pi\lambda_j t + \beta_j \sin 2\pi\lambda_j t); \; \lambda_j = j/12; \; \beta_6 \equiv 0 \qquad (2.1)$$

où les $\lambda_j = 1,2,3,\ldots,6$ sont les fréquences saisonnières et les périodes correspondantes d leur cycle (en mois) sont égales à 12, 6, 4, 3, 2.4 et 2 respectivement.

Le modèle (2.1) est dit déterministe si α_j et β_j sont constants, et stochastique si α_j e β_j sont des variables purement aléatoires sans corrélation mutuelle.

Si une série suit le modèle (2.1), son spectre est nul sauf pour six lignes spectrales au fréquences saisonnières λ_j. Cependant, pour la plupart des séries chronologiques économi ques, à saisonnalité, au lieu d'être parfaitement périodique, se modifie graduellement. Dans ces cas, le modèle (2.1) peut s'écrire

$$S_t = \sum_{j=1}^{6} (\alpha_{jt}\cos 2\pi\lambda_j t + \beta_{jt}\sin 2\pi\lambda_j t); \; \lambda_j = j/12; \; \beta_{6t} \equiv 0 \tag{2.2}$$

ù α_{jt} et β_{jt} sont des paramètres se modifiant lentement. α_{jt} et β_{jt} peuvent être soit des onctions déterministes du temps, soit des processus stochastiques dans les spectres des- uels les composantes à basse fréquence jouent un rôle prédominant.

Les séries qui suivent le modèle (2.2) ont des spectres caractérisés par des sommets aux réquences saisonnières λ_j. La puissance du spectre est nulle hors des bandes de fréquences aisonnières $\lambda_s(s)$ définies par

$$\lambda_s(\delta) = \left\{\lambda_s \text{ in } (\lambda_j - \delta, \lambda_j + \delta), \; j = 1,2,\ldots,5, (\lambda_6 - \delta, \pi)\right\} \tag{2.3}$$

où la valeur de S dépend du rythme auquel λ_{jt} et β_{jt} se modifient.

Plus les sommets sont étroits, plus les variations saisonnières sont régulières. Si l'on estime les variations saisonnières par des méthodes de moyenne mobile, plus cette dernière est courte, plus l'estimation saisonnière peut varier vite et plus la bande de fréquences relative à chaque fréquence saisonnière est censée être large.

Le spectre de puissance des fréquences inférieures à $\lambda = \frac{1}{18}$ est généralement attribué aux variations du cycle conjoncturel et la puissance des fréquences restantes qui sortent des bandes saisonnières, aux fluctuations irrégulières.

Les spectres des trois séries analysées ici ont été estimés après élimination de l'effet d'une tendance exponentielle. Si l'on examine le spectre estimé de l'IPC de tous les articles à la figure 2.2 on remarque qu'il n'y a aucune puissance à la fréquence saisonnière fondamentale $\lambda_1 = \frac{1}{12}$ et que le sommet saisonnier dominant se produit à $\lambda_3 = \frac{1}{4}$, ce qui correspond à un cycle de quatre mois. Les deux autres sommets les plus importants se produisent à λ_5 et λ_6 , qui correspondent à des cycles de 2.4 et de deux mois respectivement. La majorité des séries temporelles économiques se caractérisent par un genre opposé de spectre, présentant des sommets plus élevés à des bandes saisonnières plus faibles, autour de la

fondamentale et de sa première harmonique (voir par exemple Nerlove, Grether et Carvalho, 1979).

Une autre caractéristique importante est la largeur de ces bandes saisonnières, indiquant que le caractère saisonnier change rapidement.

L'IPC de tous les articles est un ensemble de 400 groupes élémentaires de produits. Ces derniers peuvent être regroupés en deux grandes catégories, l'Alimentation et l'Ensemble des articles hors alimentation, qui ont des causes saisonnières différentes. Comme il a été indiqué précédemment, nous avons aussi estimé le spectre de ces deux séries.

La figure 2.3 montre le spectre de l'Indice des prix de l'alimentation, qui a un profil analogue à celui des principaux indicateurs économiques.

Il convient également de noter une puissance élevée aux faibles fréquences correspondant aux cycles de trois et quatre ans, probablement en raison des cycles de production du boeuf et du porc.

La figure 2.4 montre le spectre de l'Indice des prix de tous les articles hors alimentation. Contrairement au cas précédent, le spectre a une plus grande puissance aux fréquences saisonnières plus élevées. Deux sommets apparaissent à λ_2 et λ_6, correspondant aux cycles de six et deux mois, et l'on observe aussi une puissance marquée à λ_3, λ_4 et λ_5 alors qu'il n'y en a aucune à la bande de fréquence saisonnière fondamentale. Pour nombre des articles inclus dans cette série, les prix sont établis avec une périodicité différente dépassant un mois et, en fait, cette fréquence variable du relevé des prix introduit la saisonnalité artificielle que révèle le spectre de cette série. De plus, les bandes saisonnières sont plus larges que celles des séries de l'indice des prix de l'alimentation, ce qui implique un caractère saisonnier en évolution rapide.

3. Désaisonnalisation des indices de prix à la consommation

La désaisonnalisation des séries temporelles économiques pose généralement le problème du choix de la méthode à appliquer et la question de savoir si l'on doit employer un facteur saisonnier concourant ou prévu pour l'observation courante. De plus, si la série en ques-

(All Items) Month to Month Percentage Change

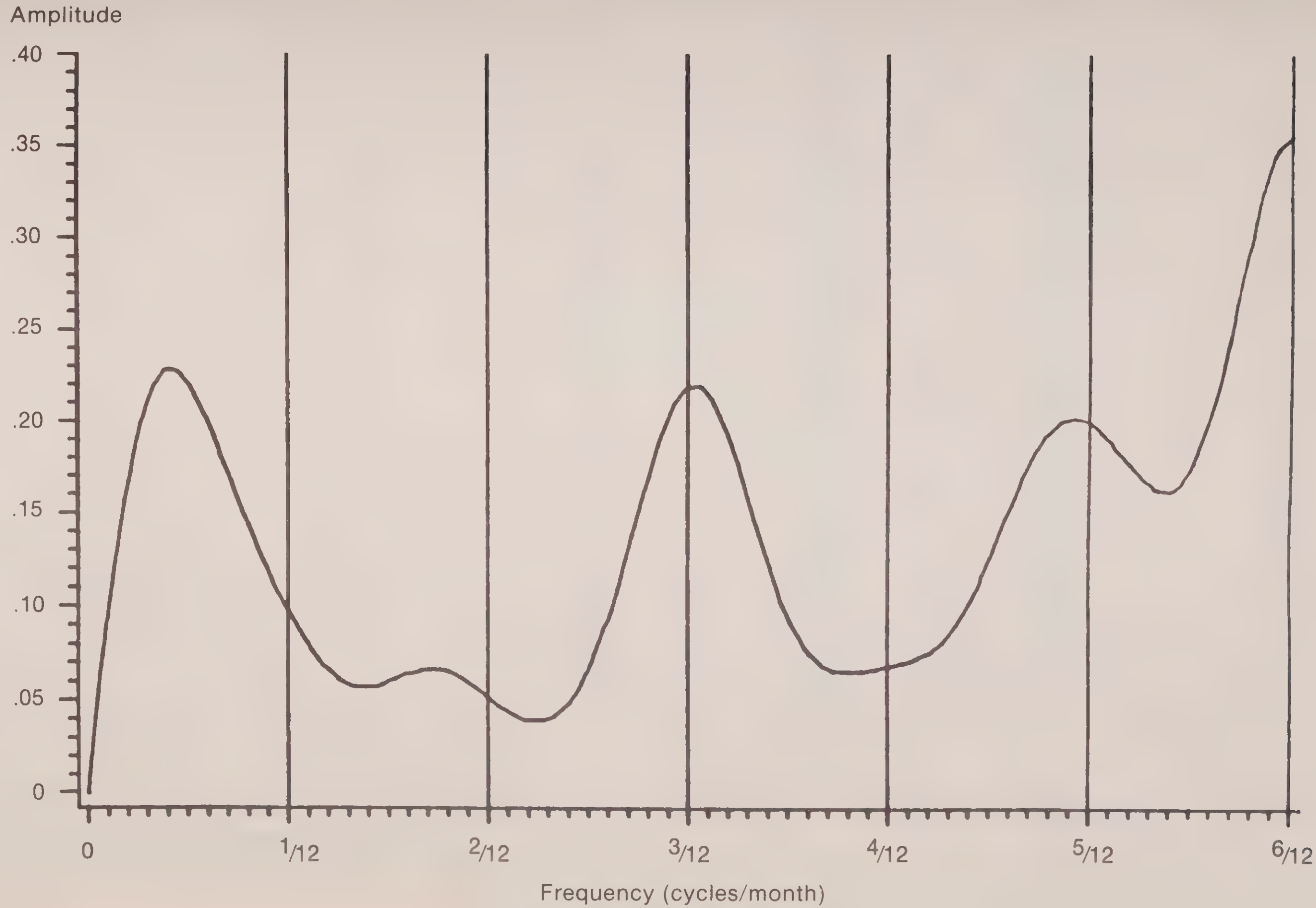
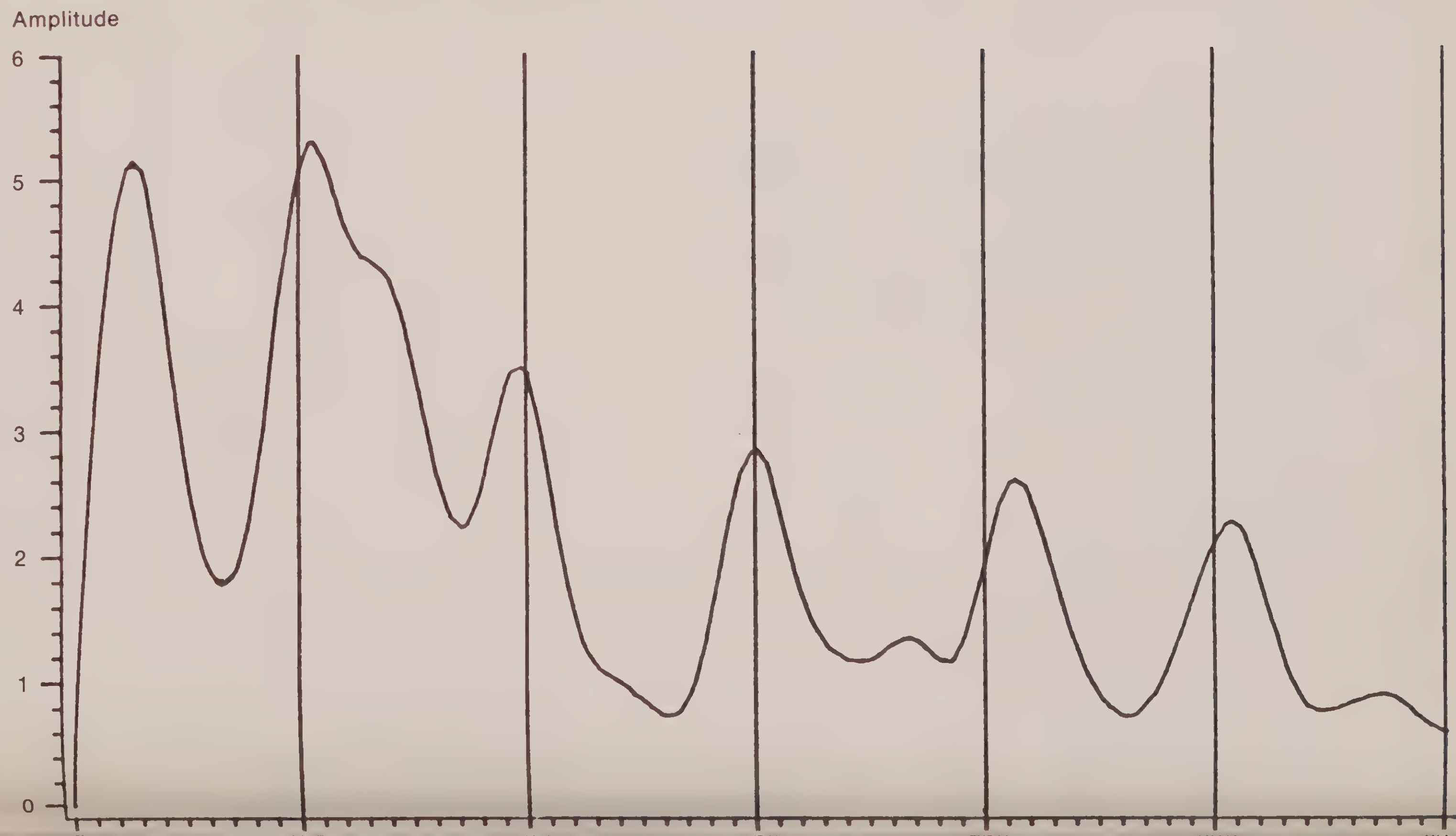
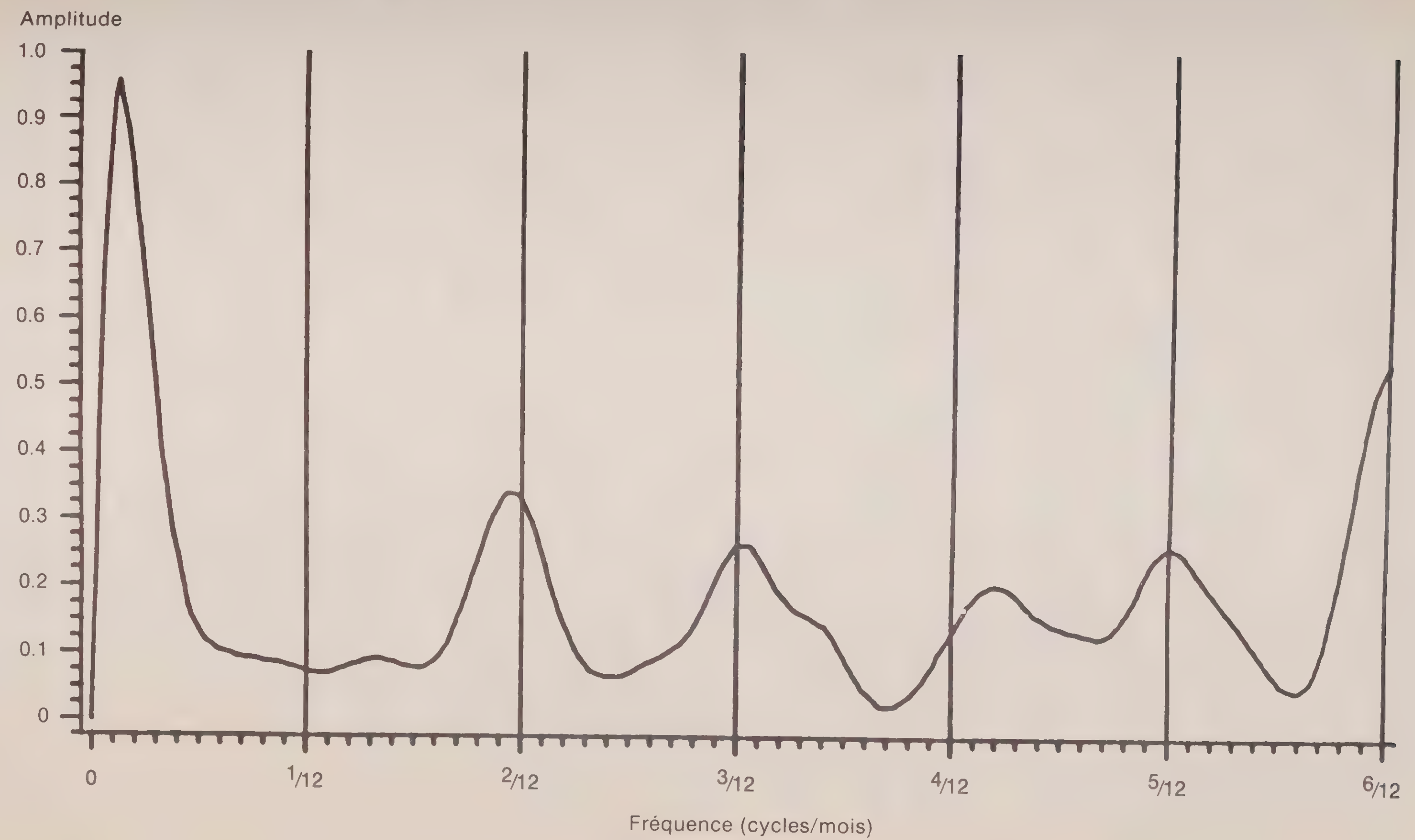


Figure 2.3

Estimations du spectre de puissance de l'IPC non corrigé (alimentation) — Pourcentages de variation mensuelle



(tous articles hors alimentation) — Pourcentages de variation mensuelle



tion est agrégative, comme l'indice des prix à la consommation, il faut décider du niveau d'agrégation auquel la désaisonnalisation doit être effectuée.

Le choix de la "meilleure" méthode dans chaque situation dépend d'un ensemble de critères qui sont déterminés principalement par le but de la désaisonnalisation. Dans le cas de l'Indice des prix à la cosommation, les critères choisis ici sont l'ampleur des révisions et la régularité des données désaisonnalisées.

3.1 Méthodes de désaisonnalisation

La plupart des méthodes de désaisonnalisation élaborées jusqu'ici sont basées sur des modèles temporels à une variable où l'estimation des facteurs saisonniers se fait de manière simple et mécanique, et non à partir d'une explication causale du phénomène étudié. Il y a eu très peu de tentatives dans ce dernier domaine et aucune d'entre elles n'est allée plus loin que le stade expérimental.

Par contre, les méthodes à une variable de désaisonnalisation des séries temporelles visent à estimer le mécanisme générateur des observations, à partir du postulat simple voulant que la série se compose d'un élément systématique et d'un élément aléatoire. La faisabilité de cette décomposition a été prouvée par Wold (1938) à l'aide d'un théorème selon lequel, si une série chronologique X_t et stationnaire jusqu'au second degré (c'est-à-dire que sa moyenne ne dépend pas du temps t et que son autocovariance ne dépend que du retard), elle peut être représentée univoquement comme la somme de deux processus sans corrélation mutuelle, l'un étant un processus déterministe linéaire ξ_t et l'autre un processus de moyenne mobile η_t infinie. On a donc

$$X_t = \xi_t + \eta_t; \ \eta_t = \sum_{j=0}^{\infty} \beta_j U_{t-j}; \ \Sigma\beta_j^2 < \infty \tag{3.2}$$

où

$$E(U_t U_s) = \sigma^2 \qquad t = s$$

$$= 0 \qquad t \neq s$$

$$E(\eta_t \xi_t) = 0 \qquad \text{pour tous les } t,s$$

t ξ_t est déterministe linéaire. Cette décomposition reste valable pour les séries non sta-ionnaires comme les séries chronologiques économiques si elles peuvent être rendues sta-ionnaires par des transformations convenables.

Dans la grande famille des modèles temporels à une variable, on peut distinguer deux atégories principales, les modèles à régression et les modèles à moyenne mobile. Voici ne brève description des principales hypothèses sous-jacentes à ces deux genres de modèles.

La plupart des travaux relatifs aux modèles de désaisonnalisation à régression reposent ır le postulat que l'élément systématique d'une série peut être représenté avec une bonne pproximation par des fonctions simples du temps sur *toute la durée* de la série. En général, eux genres de fonctions du temps sont envisagés. L'une est un polynôme de degré relative-ent faible satisfaisant au postulat que le phénomène économique évolu lentement, gulièrement et progressivement dans le temps (la tendance). L'autre est une combinaison néaire de sinus et de cosinus de fréquences différentes représentant des oscillations, stricte-ent périodiques ou non, qui influent également sur la variation totale de la série (le cycle l'élément saisonnier).

Ces fonctions sont estimées par des méthodes des moindres carrés. Pour obtenir des timations efficaces, il faut que l'élément aléatoire soit indépendant; autrement, une ver-on appropriée des moindres carrés généralisés doit être appliquée. Si la relation entre s composantes (tendance, cycle, caractère saisonnier, éléments irréguliers) est censée être ultiplicative, la méthode usuelle consiste à se servir des logarithmes et de leurs différences our mettre le mécanisme générateur de la série sous une forme additive ayant une partie éatoire stationnaire.

D'importantes contributions à l'élaboration de ce genre de modèles à régression ont été portées par Hannan [1960], Lovell [1963], Rosenblatt [1963], Ladd [1964], Jorgenson

[1964] et Henshaw [1966].

Pour surmonter les limitations qu'entraîne l'emploi d'une représentation globale de la tendance-cycle, Duvall [1966] et Stephenson et Farr [1972] ont utilisé des polynômes locaux (fonctions spline ou quasi-continues) pour de courts segments successifs de la série. Ces modèles à régression impliquaient cependant un comportement déterministe des composantes de la série. Plus tard, Pierce [1978], Gersovitz et MacKinnon [1978] et Havenner et Swamy [1978] ont pris en compte la possibilité d'un comportement stochastique de la tendance-cycle et des composantes saisonnières en mettant au point des modèles mixtes ou des modèles à régression ayant des paramètres évolutifs.

Les modèles à régression ont rarement été utilisés par les organisations statistiques officiels pour désaisonnaliser leurs données. La principale raison de cet état de choses est que, jusqu'à une date très récente, les méthodes mises au point supposaient un comportement déterministe des composantes et qu'elles n'ont pas encore réglé le problème tenant à ce que les série désaisonnalisées soient entièrement révisées lorsque de nouvelles observations viennent s'ajouter; les valeurs estimées récentes sont fortement influencées par des observations éloignées.

La majorité des méthodes de désaisonnalisation appliquées par les organisations statistiques appartiennent à la catégorie des modèles à moyenne mobile, lesquelles reposent sur le postulat que, bien que le signal d'une série temporelle soit une fonction régulière, on ne peut en donner une bonne approximation par des fonctions mathématiques simples sur toute sa durée. Étant donné que les moyennes mobiles sont des transformations linéaires, elles possèdent les propriétés de conservation de l'échelle et d'additivité. De plus, elles possèdent la propriété d'invariance dans le temps que ne partagent pas les méthodes de régression. La propriété d'invariance dans le temps signifie que, si deux éléments X_t et $X_{t+\tau}$ de la moyenne mobile L sont identiques, au déplacement dans le temps près, les produits LX_t et $LX_{t+\tau}$ sont également les mêmes, au déplacement dans le temps près. Autrement dit, la moyenne mobile ou le filtre linéaire L réagit toujours de la même manière.

Les méthodes basées sur les techniques de moyenne mobile supposent que la tendance-cycle et les composantes saisonnières changent dans le temps de manière stochastique. Les méthodes de désaisonnalisation qui appartiennent à cette catégorie sont des méthodes non

paramétriques principalement descriptives en ce sens qu'elles manquent de modèles paramétriques explicites pour chaque composante non observée.

Au cours des dernières années, cependant, plusieurs tentatives ont été faites afin d'élaborer des méthodes faisant appel à un modèle, dans lesquelles on suppose explicitement l'existence de modèles statistiques à une variable pour chaque composante. Comme il ne s'agit pas de modèles de nature causale ou explicative, la stabilité de leurs paramètres estimés est sérieusement compromise lorsque les séries sont influencées par des événements exogènes.

Les modèles explicites relèvent principalement du type gaussien ARMMI (autorégressif à moyenne mobile intégrée) mis au point par Box et Jenkins [1970] ou de variantes de ce modèle (par exemple Burman, 1980; Hillmer et Tiao, 1982; et Nerlove, Grether et Carvalho, 1979). D'autres genres de modèles (non ARMMI) ont été proposés par Akaike et Ishiguro [1980] et Bilongo et Carbone [1981].

Ces nouvelles méthodes en sont encore à l'étape du développement, et la plupart des méthodes de moyenne mobile adoptées officiellement par les organismes statistiques sont du genre non paramétrique (voir Kuiper, 1978). Parmi elles, la variante X-11 de la Method II mise au point par Shiskin, Young et Musgrave [1967] et la méthode X-11-ARMMI mise au point par Dagum [1975 et 1980] sont les plus largement appliquées. La X-11-ARMMI est une version modifiée de la variante X-11 qui consiste essentiellement à ajouter, à chaque extrémité de la série initiale, des valeurs extrapolées fournies par des modèles ARMMI et à désaisonnaliser ensuite la série prolongée. Le nouvel ensemble de coefficients de moyenne mobile résulte de la combination des filtres saisonniers X-11 et des filtres ARMMI d'extrapolation. Les filtres de désaisonnalisation de X-11 et de X-11-ARMMI diffèrent aussi bien dans leur application à la dernière observation disponible qu'à l'égard des données des dernières années. Seul le filtre symétrique appliqué aux observations centrales est le même dans les deux cas. Si l'on n'applique pas l'option ARMMI, le X-11-ARMMI se ramène à la méthode X-11.

Les révisions apportées aux valeurs désaisonnalisées courantes grâce à la méthode X-11-ARMMI avec ou sans l'option ARMMI (X-11), ainsi que pour toute méthode basée sur des modèles à moyenne mobile, sont dues aux facteurs suivants: (1) les différences des

moyennes mobiles ou des filtres linéaires de lissage appliquées à la même observation lorsque de nouvelles données deviennent disponibles; et (2) les innovations qui sont introduites dans la série avec de nouvelles observations.

À la différence des méthodes de désaisonnalisation faisant appel à la régression, les révisions de données désaisonnalisées s'arrêtent lorsque les filtres symétriques sont appliqués. Dans le contexte des deux méthodes étudiées ici, cela signifie que la première et les 3-1/2 dernières années d'une série donnée seront révisées parce qu'une bonne approximation de leur filtre exige sept années de données pour produire une estimation centrale (voir Young, 1968 et Wallis, 1974). La révision totale due aux différences entre les filtres non centraux et centraux appliqués à la même observation lorsque sa position change par rapport à la fin de la série a été mesurée dans le cas des deux méthodes par l'un des auteurs (Dagum, 1982.a et 1982.b). Les résultats obtenus indiquaient clairement que l'utilisation de valeurs extrapolées fournies par des modèles ARMMI réduisait sensiblement les mesures de la révision des filtres. Les mesures de révision étaient beaucoup plus faibles pour les bandes de fréquences saisonnières plus basses (principalement autour de la fréquence saisonnière fondamentale) que pour celles plus élevées.

Les conclusions tirées de cette étude théorique sont conformes aux résultats données dans plusieurs travaux empiriques (Dagum, 1975 et 1978, Farley et Zeller, 1978; Kuiper, 1978 et 1981, Pierce, 1980 et Kenny et Durbin, 1982).

Si l'on suppose que les valeurs extrêmes ont été éliminées et que la série est décomposée additivement, les révisions totales de la série désaisonnalisée réelle dépendent de l'ampleur de la révision du filtre multipliée par le spectre de puissance de la série; plus le spectre de puissance de la série de départ dans les fréquences saisonnières est élevé, plus la révision totale est importante. Par conséquent, les séries s'accompagnant de faibles variations saisonnières ou de profils saisonniers très réguliers peuvent être aussi bien désaisonnalisées par la méthode X-11-ARMMI avec ou sans les valeurs extrapolées.

La méthode de désaisonnalisation adoptée par Statistique Canada est la méthode X-11-ARMMI. Cependant, comme la série de l'Indice des prix à la consommation analysée dans cette étude a des variations saisonnières présentant des caractéristiques spéciales, des

tests ont été effectués afin d'évaluer si l'utilisation de l'option ARMMI améliorait ou non l'exactitude de la série désaisonnalisée courante.

3.2 Facteurs saisonniers concourants ou prévus

Des renseignements fournis par les séries désaisonnalisées jouent un rôle crucial dans l'analyse de la conjoncture économique actuelle, en particulier pour déterminer l'étape du cycle à laquelle se trouve l'économie. Cette connaissance est utile pour la prévision des mouvements cycliques ultérieurs et fournit la base des décisions de régulation économique.

Un usage important des indices désaisonnalisés de prix à la consommation est le calcul des taux de croissance de ces séries. Le gouvernement, après avoir fixé des objectifs d'augmentation afin d'orienter la politique publique, compare à ces objectifs les taux effectifs d'augmentation après élimination des variations saisonnières. La correspondance entre les taux effectifs et les taux visés dépend en partie de la qualité de la désaisonnalisation. Des erreurs dans la correction des variations saisonnières se traduisent par des taux désaisonnalisés d'augmentation erronés, ce qui entraîne des incertitudes sur le degré effectif de réalisation des objectifs. Aussi les révisions apportées aux séries désaisonnalisées courantes ont-elles toujours préoccupé les responsables de la politique, en particulier si les révisions sont fortes ou qu'elles modifient le sens des mouvements conjoncturels.

On peut obtenir une valeur désaisonnalisée **courante** en appliquant soit un facteur saisonnier prévu, soit un facteur saisonnier concourant. Ce dernier est obtenu en désaisonnalisant chaque mois toutes les données disponibles jusqu'à et y compris celles du mois en question. Le premier facteur est obtenu par la désaisonnalisation d'une série terminée un an auparavant.

Pour les séries d'indices de prix à la consommation, l'usage actuel consiste à appliquer la méthode de désaisonnalisation une fois par an aux données se terminant en décembre de cette année-là (année t); les facteurs prévus ou projetés pour l'année a venir (année t + 1) servent alors à désaisonnaliser les données à mesure qu'elles deviennent connues.

Le recours à des facteurs saisonniers prévus est habituellement justifié par le désir d'accroître la confiance du public à l'égard des grands indicateurs économiques désaisonnalisés,

étant donné que ces facteurs sont publiés avant les mois auxquels ils s'appliquent. Cependant, l'estimation des facteurs saisonniers projetés ne tient pas compte de l'évolution toute récente de la série considérée, comme lorsqu'on utilise des facteurs saisonniers concourants.

Les travaux empiriques menés à Statistique Canada ont conduit à utiliser des facteurs concourants pour la désaisonnalisation des statistiques de population active pour la première fois en 1975, à l'aide de la méthode X-11-ARMMI. Les études empiriques plus récentes ont montré l'avantage que présentait l'utilisation de facteurs saisonniers concourants au lieu de facteurs prévus, sous l'angle des révisions (voir par exemple Bayer et Wilcox, 1981; Cleveland, Grambsch et Terpenning, 1982; Kenny et Durbin, 1982 et McKenzie, 1982). D'autres éléments théoriques militant en faveur d'une désaisonnalisation concourante sont fournis dans Dagum [1983.c] pour les méthodes X-11 et X-11 ARMMI. Cependant, l'avantage que donne l'utilisation d'un facteur saisonnier concourant est pratiquement nul si le caractère saisonnier est extrêmement régulier ou si les valeurs les plus récentes, sont fortement influencées par les valeurs extrêmes.

Pour les séries analysées ici, des tests ont été effectués afin d'évaluer celui des deux facteurs qui était préférable pour leur désaisonnalisation courante.

3.3 Problème du niveau d'agrégation

L'Indice des prix à la consommation est une mesure de variations des prix de détail payés par la population cible pour les biens et services de consommation; à ce titre, il constitue une moyenne pondérée des rapports de prix relatifs à tous les produits particuliers faisant partie du "panier". Comme l'IPC est un agrégat, la question classique se pose à propos de sa désaisonnalisation: est-il préférable de désaisonnaliser d'abord les rapports de prix relatifs aux produits particuliers (le cas échéant) avant de les agréger à l'aide des coefficients de pondération, ou une désaisonnalisation de l'indice agrégatif des prix produit-elle de meilleurs résultats? La première méthode est appelée dans la litérature correction indirecte, tandis que la dernière méthode est appelée désaisonnalisation directe de données agrégatives.

Pour compliquer encore le problème, il n'est pas nécessaire dans la méthode indirecte de corriger les variations saisonnières au niveau des rapports de prix de chaque produit.

Il est possible de construire l'IPC désaisonnalisé relatif à tous les articles en partant d'indices agrégatifs de niveau inférieur, par exemple d'indices de prix relatifs à de grandes catégories de produits, tant que les coefficients de pondération appropriés sont utilisés.

Bien entendu, le choix de la méthode la plus appropriée ne poserait pas de problème si les estimations désaisonnalisées fournies par les diverses techniques étaient identiques ou au moins statistiquement équivalentes, mais cela n'est généralement pas le cas. Comme l'ont souligné Lothian et Morry [1977], les méthodes de désaisonnalisation les plus souvent employées, la variante X-11 de la Method II et la X-11-ARMMI, présentent des caractères non linéaires qui introduisent des différences entre les estimations produites par la méthode directe et la méthode indirecte. Les caractères non linéaires de ces deux méthodes peuvent être éliminés si les conditions suivantes sont respectées:

(1) toutes les séries composantes ainsi que la série agrégative sont désaisonnalisées à l'aide du modèle à décomposition additive;

(2) les valeurs extrêmes ne sont pas remplacées;

(3) la courbe variable tendance-cycle (filtre d'Henderson) est maintenue constante pendant toutes les itérations.

Le respect de ces conditions est toutefois généralement obtenu au détriment de la qualité de la désaisonnalisation.

Une autre façon d'aborder la désaisonnalisation des données agrégatives, proposée par Geweke [1978], fait appel à l'espérance mathématique conditionnelle de l'agrégat par rapport à toutes ses composantes prises conjointement. Cette méthode est un cas particulier de la théorie générale de prévision dans les processus stochastiques stationnaires mise au point par Wold [1938], Kolmogorov [1939], Wiener [1949] et Whittle [1963].

Bien que la proposition de Geweke soit séduisante sur le plan théorique, elle présente plusieurs inconvénients pratiques, dont le principal est que la méthode exige la connaissance des distributions conjointes de l'agrégat, des composantes saisonnières et des éléments non saisonniers, condition qui ne peut être satisfaite que par des séries simulées.

Dans l'exposé qui suit, le problème de la désaisonnalisation d'une série d'indice agrégati de prix sera envisagée dans le contexte de la méthode X-11-ARMMI.

Sur le plan conceptuel, ni la méthode directe ni la méthode indirecte ne sont optimales Il existe des arguments en faveur de chacune.

La méthode directe:

(i) a l'avantage de fournir des données désaisonnalisées historiques, peu importe la fré quence des mises à jour du panier de consommation (la qualité de la désaisonnalisa tion pourrait cependant être diminuée par d'importants changements de la pondéra tion des dépenses);

(ii) est d'un fonctionnement plus rapide et moins coûteux;

(iii) présente un effet d'annulation possible à l'égard des variations mensuelles de pro duits saisonniers ayant des saisons opposées.

Par ailleurs, la méthode indirecte:

(i) fournit le moyen analytique d'attribuer le changement de l'IPC de tous les article aux variations des prix de divers produits;

(ii) permet une désaisonnalisation au niveau élémentaire lorsque la saisonnalité est bie identifiable et affiche un profil relativement simple – ces dernières propriétés ayan été étudiées en détail par Dagum [1979].

Si les méthodes directe et indirecte produisent des estimations qui ne diffèrent pas trop le choix de "la meilleure" méthode doit être fonction de l'importance relative des avan tages précédents. Par contre, si les deux ensembles de valeurs désaisonnalisées diffèren sensiblement, la préférence doit être donnée à la méthode qui produit des estimations plu exactes, d'après des critères pertinents. Pour les indices de prix à la consommation, le critères choisis étaient l'ampleur des révisions et la régularité de la série désaisonnalisée

4. Critère d'évaluation

La régularité d'une série désaisonnalisée est généralement une propriété souhaitable pour les responsables de la politique. On s'attend à ce que l'élimination de la composante saisonnière supprime une source de variabilité récurrente de la série, de sorte que cette dernière, une fois désaisonnalisée, doit être moins irrégulière que les données de départ. La régularité ne devrait cependant pas être le critère unique employé pour choisir une méthode de désaisonnalisation, car il est toujours possible d'avoir une série désaisonnalisée fortement lissée au moyen de facteurs saisonniers qui fluctuent de manière moins régulière. En fait, en ajustant les facteurs saisonniers suffisamment bien aux données, on peut obtenir à peu près n'importe quel degré de régularité des séries désaisonnalisées. Pour éviter cela, il faut envisager un second critère, l'ampleur des révisions des facteurs saisonniers. La minimisation des révisions apportées aux données désaisonnalisées fait l'objet de beaucoup d'attention actuellement. Les révisions se produisent, dans la méthode actuelle de désaisonnalisation, X-11-ARMMI, en raison des différences des filtres asymétriques appliqués aux années les plus récentes de données et aux innovations qui entrent dans la série avec de nouvelles observations.

Il est important, pour l'élaboration de la politique économique, que les révisions apportées aux statistiques désaisonnalisées courantes soient les plus faibles possible. Comme une valeur désaisonnalisée courante peut être obtenue en appliquant un facteur saisonnier concourant ou projeté, les mesures des révisions sont définies ci-après pour ces deux cas.

4.1 Mesure des révisions

Comme il était indiqué à la section 3.1, les révisions des séries désaisonnalisées par les méthodes X-11 et X-11-ARMMI cessent lorsque les filtres symétriques centraux sont appliqués. Dans le contexte de ces deux méthodes, cela signifie que la première et les 3-1/2 dernières années seront révisées parce qu'une bonne approximation de leurs filtres nécessite 85 points de données mensuelles. La révision totale d'un facteur saisonnier *concourant* $\hat{S}_t^c$ est définie par

$$R(\hat{S}_t^c) = \hat{S}_t^c - \hat{S}_t^f \qquad (4.1.1)$$

Où $\hat{S}_t^f$ est le facteur saisonnier ''final'', en ce sens qu'il ne changera plus lorsque de nouvelles données s'ajouteront à la série. De même, pour la série désaisonnalisée concourante notée $C\hat{I}_t^c$, la mesure de la révision totale est[1]

$$R(C\hat{I}_t^c) = C\hat{I}_t^c - C\hat{I}_t^f \tag{4.1.2}$$

En remplaçant $\hat{S}_t^c$ par le facteur saisonnier projeté $\hat{S}_t^p$ dans l'équation (4.1.1), nous obtenons une mesure de la révision pour le facteur saisonnier projeté, et de même pour la série désaisonnalisée courante dans l'équation (4.1.2), si un facteur saisonnier projeté est appliqué.

Dans le contexte de la méthode X-11-ARMMI, au moins cinq années de données sont nécessaires pour produire une série désaisonnalisée; nous aurions alors besoin de 9-1/2 années pour obtenir douze points t pour lesquels les révisions totales définies précédemment puissent être calculées. Comme la longueur de la série analysée dans cette étude est de huit ans, de janvier 1974 à décembre 1981, les équations (4.1.1) et (4.1.2) ne peuvent s'appliquer (il n'y a pas de facteur saisonnier ''final''). Pour surmonter cette limitation due au manque de points de données, la méthode suivie ici consiste à comparer chacun des facteurs saisonniers concourants et projetés pour les trois dernières années aux estimations correspondantes obtenues à partir de la désaisonnalisation des données jusqu'à la fin de l'année 1981, notées $\hat{S}_t^{81}$. Ainsi, les mesures modifiées de la révision pour les facteurs saisonniers concourants et projetés sont

$$R_1(\hat{S}_t^c) = \hat{S}_t^c - \hat{S}_t^{81} \tag{4.1.3}$$

$$R_1(\hat{S}_t^p) = \hat{S}_t^p - \hat{S}_t^{81} \tag{4.1.4}$$

et de même, pour la série désaisonnalisée courante,

$$R_1(C\hat{I}_t^c) = C\hat{I}_t^c - C\hat{I}_t^{81} \tag{4.1.5}$$

$$R_1(C\hat{I}_t^p) = C\hat{I}_t^p - C\hat{I}_t^{81} \tag{4.1.6}$$

Les substitutions précédentes sont justifiées, car les filtres saisonniers concourants et prévus convergent de façon monotone vers le filtre central. L'ampleur réelle des révisions totales sera toutefois plus importante qu'avec les mesures modifiées.

Les mesures de révision calculées pour la série analysée dans cette étude sont les révisions absolues moyennes de la série désaisonnalisée courante obtenue à l'aide de facteurs saisonniers concourants (c) et projetés (p), soit

$$|\bar{R}^{c}| = \sum_{t=61}^{96} |C\hat{I}_t^{c} - C\hat{I}_t^{81}| \;/36 \qquad (4.1.7)$$

$$|\bar{R}^{p}| = \sum_{t=61}^{96} |C\hat{I}_t^{p} - C\hat{I}_t^{81}| \;/36 \qquad (4.1.8)$$

La méthode donnant la plus faible révision absolue moyenne est celle qui doit être choisie.

.2 Mesures de la régularité

Dans le contexte des séries d'Indice de prix à la consommation, le niveau est moins important que le pourcentage de variation mensuelle, aussi appelé taux de croissance. Ainsi, régularité des données désaisonnalisées sera mesurée par leur taux de croissance r_t.

La mesure utilisée est la suivante

$$S = \sqrt{\sum_{t=62}^{96} (r_t - \bar{r})^2 \;/35} \qquad (4.2.1)$$

où $$r_t = \frac{C\hat{I}_t - C\hat{I}_{t-1}}{C\hat{I}_{t-1}} \times 100 \qquad (4.2.2)$$

et $$\bar{r} = \sum_{t=62}^{96} r_t \;/35 \qquad (4.2.3)$$

S est une mesure de la dispersion autour du taux moyen de croissance au cours des trois dernières années. Les estimations désaisonnalisées de l'équation (4.2.2) peuvent s'appliquer aux estimations concourantes, projetées ou de décembre 1981. Lorsque les méthodes de désaisonnalisation produisent des estimations présentant la même valeur de S, celle-ci est complétée par la mesure D, visant à détecter la persistance du taux de croissance en niveau et définie par

$$D = \sum_{t=63}^{96} | r_t - r_{t-1} | \;/34 \qquad (4.2.4)$$

La préférence sera alors donnée à la méthode ayant la plus faible valeur de D.

4.3 Révisions des facteurs saisonniers et des taux de croissance

Il est possible de mesurer l'effet que les révisions des estimations de facteurs saisonniers ont sur les taux de croissance mensuelle des indices de prix (voir Maravall, 1981). Si l'on suppose qu'un retard d'un mois est une unité relativement petite, l'équation (4.2.2) peut s'écrire

$$r_t = \frac{d\,C\hat{I}_t}{C\hat{I}_t} = d \log C\hat{I}_t \simeq \log C\hat{I}_t - \log C\hat{I}_{t-1} \qquad (4.3.1)$$

où d symbolise l'opérateur différentiel.

La série, désaisonnalisée CI_t est égale à la série de départ X_t divisée par le facteur saisonnier estimatif S_t, c'est-à-dire

$$C\hat{I}_t = X_t / \hat{S}_t \qquad (4.3.2)$$

Par conséquent, par substitution de l'équation (4.3.2) dans l'équation (4.3.1), l'expression devient

$$r_t \simeq \Delta\log X_t - \Delta\log \hat{S}_t \qquad (4.3.3)$$

où est l'opérateur de différence.

Si l'on note ϵ_t l'erreur de révision de S_t, de sorte que

$$\hat{S}_t = S_t + \epsilon_t \tag{4.3.4}$$

on a

$$\log \hat{S}_t = \log S_t \left(1 + \frac{\epsilon_t}{S_t}\right)$$

$$= \log S_t + \log \left(1 + \frac{\epsilon_t}{S_t}\right) \tag{4.3.5}$$

Étant donné que ϵ_t/S_t est un petit nombre, à l'aide de développement de Taylor une aproximation de l'équation (4.3.5) peut s'écrire

$$\log \hat{S}_t \simeq \log S_t + \frac{\epsilon_t}{S_t} \tag{4.3.6}$$

En substituant l'équation (4.3.6) dans l'équation (4.3.3), on a

$$r_t \simeq \Delta\log X_t - \Delta\log S_t - v_t \tag{4.3.7}$$

où

$$v_t = \frac{\epsilon_t}{S_t} - \frac{\epsilon_{t-1}}{S_{t-1}} \tag{4.3.8}$$

par conséquent, la variance de r_t due à la révision du facteur saisonnier ϵ_t est donnée par

$$\text{var}_S(r_t) = \text{var}_S\left(\frac{\epsilon_t}{S_t} - \frac{\epsilon_{t-1}}{S_{t-1}}\right) \tag{4.3.9}$$

La variance dépend de S. Si les facteurs saisonniers sont voisins de l'unité et qu'on suppose l'absence de toute erreur dans la série initiale X_t, on a

$$\text{var } r_t \simeq 2 \text{ var } \epsilon_t (1-\rho(1)) \tag{4.3.10}$$

L'équation (4.3.10) exprime que l'effet de l'erreur de facteur saisonnier sur le taux de variations mensuelles peut être calculé à partir de la variance de la série d'erreurs et de son autocorrélation de premier ordre.

5. Choix de la méthode appropriée de désaisonnalisation pour la série de l'Indice des prix à la consommation au Canada

Il y avait trois grandes questions à étudier à propos de la désaisonnalisation de la série de l'indice des prix à la consommation au Canada:

1. L'extrapolation ARMMI améliore-t-elle la qualité des estimations désaisonnalisées?

2. L'emploi de facteurs saisonniers concourants est-il préférable à l'utilisation de facteurs prévus pour produire les estimations désaisonnalisées courantes?

3. La correction indirecte est-elle meilleure que la méthode directe et, dans l'affirmative, quel est le niveau optimal de désagrégation pour la désaisonnalisation de la série de l'Indice des prix à la consommation de tous les articles?

Étant donné que l'étude des deux premiers problèmes ferait appel à des modalités extrêmement différentes selon qu'on adopterait la méthode directe ou la méthode indirecte, il fallait étudier en premier lieu le troisième problème.

5.1 Problème du niveau d'agrégation

Les données utilisées pour résoudre le problème de l'agrégation ont été fournies par la Division des prix. (Les trois dernières années figurent aux annexes 1 à 3.) Elles consistaient en trois ensembles de séries désaisonnalisées de l'IPC d'ensemble allant de janvier 1974

ı décembre 1981, 1974 étant l'année de base. La première série était la statistique désaison-ıalisée officielle de l'Indice des prix à la consommation d'ensemble, calculé comme un ıgrégat de séries d'indices de prix relatifs aux divers articles, dont certains étaient désaison-ıalisées et les autres pas. Les séries présentant une saisonnalité identifiable ont été incor-porées à l'agrégat sous forme corrigée, tandis que les statistiques ne présentant aucun pro-il saisonnier apparent étaient conservées telles quelles avant la sommation. Ce genre de néthode indirecte présente toujours le danger que les mouvements saisonniers qui passent ıaperçus dans les séries particulières à cause de la présence de fortes fluctuations irrégulières leviennent plus prononcées par la méthode d'agrégation (dans laquelle les éléments ir-éguliers s'annulent mutuellement) et produisent une saisonnalité résiduelle au niveau de 'agrégat désaisonnalisé. Les coefficients utilisés dans l'agrégation étaient fondés sur le panier e produits de 1974. (La méthode précédente coïncide avec celle utilisée pour calculer le aux publié mensuellement de septembre 1978 à décembre 1981.)

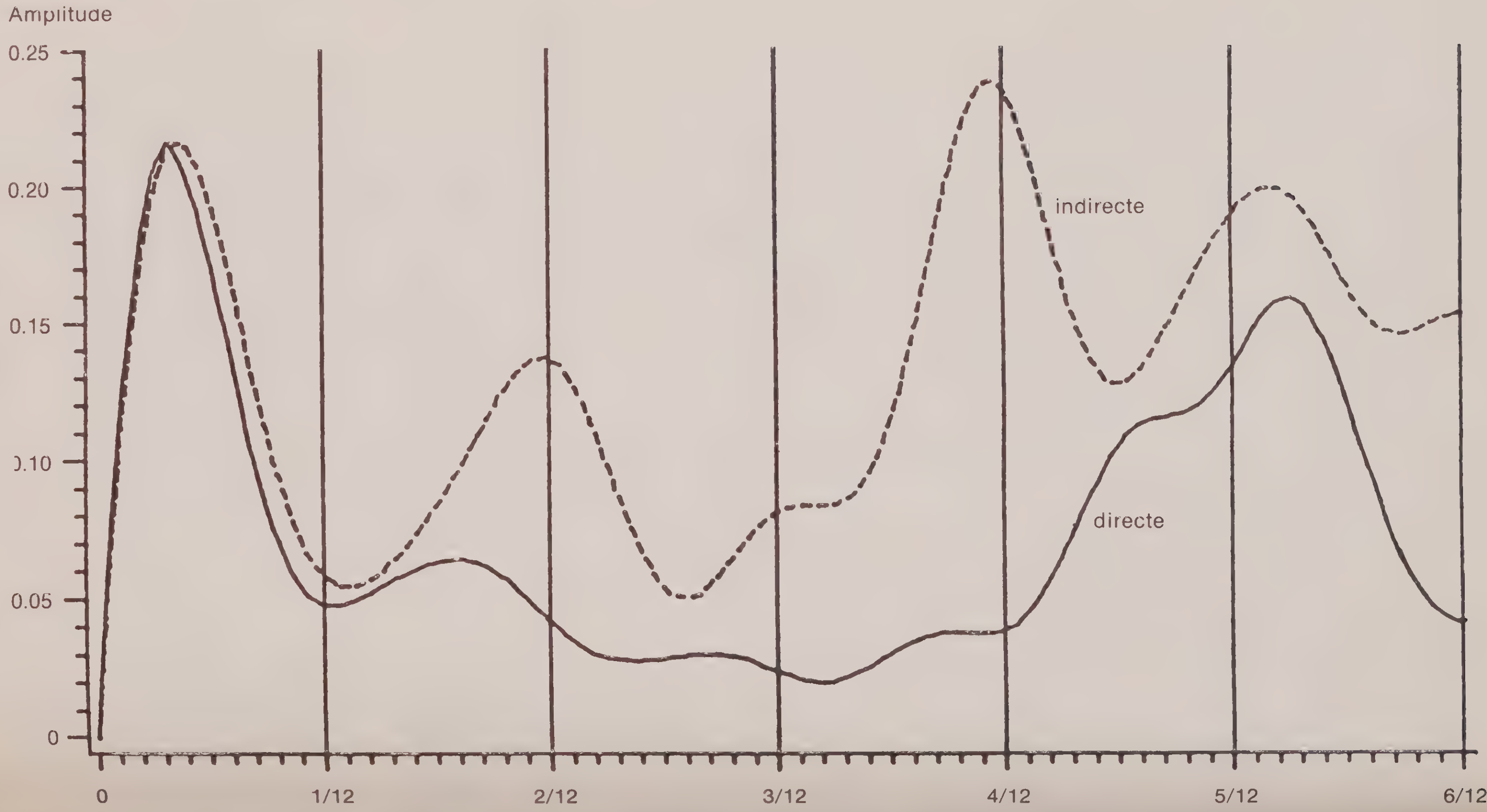
Cette statistique officielle a servi de norme de comparaison pour deux méthodes de emplacement; l'une d'elles consistait à désaisonnaliser directement la série enchaînée de IPC pour tous les articles. La seconde consistait à désaisonnaliser les sept séries enchaînées 'indices de prix relatifs aux composantes principales de produits et à en faire la somme à l'aide des profils de dépenses applicables pendant la période considérée) pour obtenir IPC désaisonnalisé de tous les articles.

Pour éviter les difficultés dues au traitement non homogène des trois méthodes en matière 'enchaînement, des tests ont été effectués sur les trois dernières années de statistiques seule-ıent. On s'assurait également ainsi que les conclusions se rapportaient aux mouvements s plus récents et, partant, les plus pertinents des séries.

Au chapitre 4, deux genres de mesures avaient été présentées pour comparer la qualité e la désaisonnalisation. Cependant, pour l'étude du problème de l'agrégation, l'analyse es trois ensembles d'estimations se limitera à l'examen du degré de régularité. Le calcul es mesures de la révision aurait mis en jeu des milliers de passages de désaisonnalisation, qui n'était pas jugé faisable sur le plan opérationnel ou financier.

Figure 5.1

Estimations du spectre de puissance de l'IPC (tous articles) désaisonnalisé — Pourcentages de variations mensuelle



Le tableau 5.1 présente un résumé des résultats. Les mesures de régularité S et D ont été calculées à partir des trois dernières années de la série désaisonnalisée $C\hat{I}_t^{81}$ obtenue par les trois méthodes différentes.

TABLEAU 5.1 Mesures de régularité de l'Indice des prix à la consommation de tous les articles pour trois méthodes de désaisonnalisation

	Données non corrigées	Correction indirecte	Correction directe	Correction indirecte (7 grands sous-groupes)
Écart-type du taux de croissance "S"	.296	.261	.190	.201
Différence absolue de taux de croissance "D"	.330	.324	.212	.226

Selon le tableau 5.1, le passage de la méthode indirecte utilisée officiellement à l'une des deux méthodes de remplacement améliore sensiblement la régularité. Alors qu'une désaisonnalisation par la méthode indirecte d'origine n'entraîne qu'une réduction de 12% des fluctuations du taux de croissance, la méthode directe les réduit de 36% (trois fois plus), comme en témoigne la valeur de 0.190, contre 0.296 dans la série brute.

Une comparaison des séries désaisonnalisées obtenues par la méthode directe et la méthode indirecte (figure 5.1) débouche sur des conclusions analogues, à savoir que la méthode directe est préférable à la méthode indirecte. Cette dernière, loin de réduire la puissance à toutes les fréquences saisonnières (comme est censée le faire une méthode de désaisonnalisation), l'accroît même aux fréquences saisonnières de 2/12 et 4/12, ce qui n'est évidemment pas acceptable.

Quand les résultats de cette recherche furent communiqués à la Division des prix, il fut décidé d'adopter la méthode directe pour la désaisonnalisation de la série de l'IPC de tous les articles à compter de la publication de l'indice d'avril 1982. La méthode directe fut préférée à la solution de rechange car elle produisait des estimations un peu plus régulières

et était également d'utilisation plus commode. Cependant, s'il se révèle nécessaire à l'avenir d'attribuer les mouvements de l'agrégat aux variations des indices de prix relatifs aux composantes principales de produits, la construction de l'IPC à partir des sept composantes désaisonnalisées est également acceptable.

En raison de l'importance particulière de deux autres séries d'indices de prix, Aliments et l'Ensemble sans les aliments, on a fait une étude de ces deux séries afin de voir s'il était souhaitable de remplacer la méthode indirecte utilisée précédemment par la méthode directe de désaisonnalisation. Dans le cas de l'IPC l'Ensemble sans les aliments, une troisième méthode a également été utilisée, l'indice désaisonnalisé étant obtenu dans ce cas comme l'agrégat de ses six sous-composantes principales désaisonnalisées.

Les tableaux 5.2 et 5.3 résument les résultats de cette dernière étude.

TABLEAU 5.2 Mesures de régularité de l'Indice des prix à la consommation – Alimentation, pour deux méthodes de désaisonnalisation

Mesure	Données non corrigées	Correction indirecte	Correction directe
Écart-type du taux de croissance "S"	.835	.770	.721
Différence absolue du taux de croissance "D"	.809	.798	.656

TABLEAU 5.3 Mesures de régularité de l'Indice des prix à la consommation – tous articles hors alimentation, pour trois méthodes de désaisonnalisation

	Données non corrigées	Correction indirecte	Correction directe	Correction indirecte (6 grands sous-groupes)
Écart-type du taux de croissance "S"	.319	.277	.255	.260
Différence absolue du taux de croissance "D"	.393	.329	.281	.273

On a obtenu les mêmes résultats que pour l'Indice des prix à la consommation de l'Ensemble. La désaisonnalisation directe produisait une statistique plus régulière que la méthode indirecte. Pour l'IPC des Aliments, les fluctuations étaient réduites de 14% par la méthode directe et de 8% par la méthode indirecte, les chiffres correspondants étant de 20% et de 13% pour l'Ensemble sans les aliments. Dans ce dernier cas, l'amélioration apportée par la méthode indirecte appliquée aux six sous-groupes était comparable à celle produite par la méthode directe, tout en étant un peu plus faible.

Au vu de ces résultats et à la suite de nos recommandations, la Division des prix a adopté la méthode directe pour les deux séries à partir de la publication de l'indice d'avril 1982. Le problème de l'agrégation étant ainsi réglé et la méthode directe adoptée, notre examen de la possibilité des extrapolations ARMMI et de la validité de facteurs saisonniers concourants était grandement simplifié.

La section 5.2 expose les résultats de l'analyse menée au sujet de ces deux questions.

5.2 Comparaison de quatre options différentes pour la production d'estimations désaisonnalisées courantes de l'Indice des prix à la consommation

Une fois la méthode directe préférée pour la désaisonnalisation de l'Indice des prix à la consommation de l'Ensemble, il reste quatre options à considérer dans le contexte de la correction directe. Ces options sont le recours à la méthode X-11-ARMMI avec et sans l'extrapolation ARMMI et, pour chaque cas, le choix de facteurs saisonniers concourants ou prévus lorsqu'on calcule les chiffres désaisonnalisés courants.

En ce qui concerne l'utilisation de l'extrapolation ARMMI, les études théoriques menées par l'un des auteurs (Dagum, 1982.a et 1982.b) ont montré que l'addition à la série brute d'une année de valeurs extrapolées réduisait sensiblement les révisions de filtres. Cependant, tout dépendant de la nature de la série à laquelle les filtres de désaisonnalisation sont appliqués, cette réduction de la distance entre les filtres peut résulter ou non en une diminution de l'ampleur des révisions que la série subira lorsque de nouvelles données deviendront connues. L'extrapolation ARMMI produit les meilleurs résultats lorque la série considérée se conforme à une certaine structure concernant le profil saisonnier et les variations irrégulières. Comme la série de l'Indice des prix à la consommation présente des structures saisonnières plutôt inhabituelles, comme l'a montré le chapitre 2, il est nécessaire

d'entreprendre une étude empirique pour déterminer s'il est profitable d'étendre la série à l'aide de valeurs extrapolées avant une désaisonnalistion.

De même, des travaux théoriques ont été menés [Dagum 1982.c] sur les propriétés des filtres saisonniers concourants et prévus de X-11-ARMMI à l'égard des révisions. Les filtres concourants sont plus voisins des filtres centraux finals que les filtres prévus, d'où des révisions moins marquées des filtres en raison de la dépendance des révisions totales à l'égard des données, en particulier en présence de valeurs extrêmes, mais des travaux empiriques sont également nécessaires pour indiquer quels facteurs produisent les meilleures estimations dans le cas de la série de l'Indice des prix à la consommation.

L'étude de la méthode appropriée (X-11-ARMMI avec ou sans extrapolation) et le genre optimal de facteurs saisonniers (concourants ou prévus) à utiliser posent un problème de cercle vicieux. Pour voir s'il est souhaitable d'utiliser les extrapolations ARMMI, il faut utiliser le genre préféré de facteurs saisonniers; par ailleurs, pour choisir entre les facteurs concourants et prévus, il faut d'abord régler la question d'une méthode convenable. Pour éviter de nous placer dans ce cercle vicieux, nous examinons les deux problèmes simultanément. Quatre ensembles de chiffres désaisonnalisés sont produits à l'aide de facteurs saisonniers concourants obtenus avec ou sans extrapolation ARMMI et de facteurs saisonniers prévus découlant d'une désaisonnalisation X-11-ARMMI avec et sans recours à l'option ARMMI.

La désaisonnalisation optimale sera donnée par la combinaison de la méthode appropriée et des facteurs saisonniers qui produisent les estimations les plus régulières avec les révisions les plus faibles.

Les données faisant l'objet de l'analyse se limitent là encore aux trois dernières années, de janvier 1979 à décembre 1981, afin de révéler les résultats les plus récents des quatre options en question sans trop réduire la taille de l'échantillon. Les mesures de régularité et de révision des quatre séries désaisonnalisées sont calculées afin de déterminer les estimations qui satisfont le mieux aux critères choisis, tels qu'ils ont été définis aux sections 4.1 et 4.2.

Le tableau 5.4 résume les résultats de l'analyse.

TABLEAU 5.4 Mesures de régularité et de révision de l'Indice des prix à la consommation de l'Ensemble pour quatre méthodes de désaisonnalisation

Mesure	Données non corrigées	X-11-ARMMI sans extrapolation		X-11-ARMMI avec extrapolation	
		Concourant	Prévu	Concourant	Prévu
S	.296	.272	.271	.237	.269
D	.330	.321	.338	.232	.300
R	–	.114	.131	.103	.125

Il ressort clairement du tableau 5.4 que, parmi les quatre options, c'est l'extrapolation ARMMI et l'utilisation de facteurs saisonniers concourants qui produisent les meilleures estimations désaisonnalisées, tant du point de vue de la régularité que de celui de l'ampleur des révisions. L'adoption de ces options combinées pour la désaisonnalisation entraîne une réduction de 20% des fluctuations du taux de croissance, contre 9% seulement avec la méthode utilisée officiellement (facteurs saisonniers prévus et extrapolation ARMMI). Ainsi, les estimations produites de cette façon donnent des mouvements mensuels plus exacts que les chiffres officiels.

L'amélioration est également sensible au niveau des révisions: la révision moyenne de 0.125 observée avec la méthode officielle est ramenée à 0.103 quand les estimations sont produites par la combinaison proposée. Il faut de plus se rappeler que, comme il a été mentionné au chapitre 4, ces révisions ne sont pas les révisions définitives et les valeurs du tableau 5.4 sous-estiment l'ampleur des révisions totales. Pour les révisions finales, la différence absolue entre les deux mesures serait beaucoup plus marquée. Si l'on n'examine que les données de l'année 1979, pour laquelle les révisions calculées sont presque identiques aux révisions finales, les mesures de révision sont de 0.167 et de 0.100 pour les deux corrections considérées.

Le passage à des facteurs saisonniers concourants pose un dilemme pour la publication du pourcentage de variation mensuelle de l'Indice des prix à la consommation. La variation mensuelle doit-elle être calculée à partir du niveau désaisonnalisé publié le mois précédent ou le niveau révisé du dernier mois doit-il servir de base au calcul du taux de variation? Cette question ne se présente pas dans le cas des facteurs saisonniers prévus, puisque

les séries sont traitées par le programme X-11-ARMMI une fois par an seulement et que, à partir de là, les données désaisonnalisées sont produites à l'aide des facteurs projetés, qui restent inchangés pendant toute l'année suivante.

Pour résoudre ce problème, nous avons calculé à nouveau les mesures de régularité et de révision introduites précédemment. Au tableau 5.4, les mesures de régularité et de révision du taux de croissance sont calculées à l'aide des **valeurs non révisées du mois précédent**. Le tableau 5.5 indique les mesures obtenues quand le taux de croissance est basé sur les **premières estimations révisées du dernier mois**.

TABLEAU 5.5 Comparaison des estimations provisoires et des premières estimations révisées du taux de croissance de l'Indice des prix à la consommation de l'ensemble à l'aide de facteurs concourants avec extrapolation ARMMI

Mesure	Estimations provisoires	Premières estimations révisées
S	.237	.200
D	232	.176
$\bar{R}$	103	.091

Il ressort du tableau 5.5 que l'incorporation des premières estimations révisées dans le calcul des pourcentages de variation mensuelle entraîne une nette amélioration sur le plan tant de la régularité que de l'ampleur des révisions.

Les résultats de l'étude empirique permettent de suggérer fortement que:

1. on obtienne l'Indice désaisonnalisé des prix à la consommation au moyen de la méthode directe ou, au moins, par la méthode indirecte appliquée aux sept composantes principales;

2. on utilise la méthode X-11-ARMMI avec l'option d'extrapolation ARMMI;

3. on calcule les estimations désaisonnalisées courantes à l'aide de facteurs saisonniers concourants plutôt que de prévisions des facteurs saisonniers;

4. le calcul des pourcentages de variation mensuelle soit basé sur les premières estimations révisées du mois précédent.

L'application de ces propositions se traduira par une série désaisonnalisée de l'IPC nettement plus régulière et qui subira des révisions plus faibles lorsque de nouvelles observations seront incorporées à la série.

6. Conclusions

Nous avons étudié dans ce document les principales caractéristiques des variations saisonnières de la série de l'Indice des prix à la consommation et avons comparé plusieurs méthodes de désaisonnalisation permettant de les estimer.

Parmi les séries considérées, le spectre de l'IPC de tous les articles a affiché des variations saisonnières qui diffèrent sensiblement de celles de la plupart des autres indicateurs économiques. Il n'y a aucune puissance à la bande saisonnière fondamentale associée au cycle annuel et les sommets principaux correspondent aux cycles de 4, 2.4 et 2 mois. Ces sommets saisonniers résultent du fait que la plupart des articles incorporés à l'indice avec de forts coefficients ont des prix recueillis avec une fréquence différente et variant dans le temps. La propriété de variation dans le temps introduit une structure saisonnière en évolution rapide qui se reflète dans le spectre par des bandes saisonnières très larges.

La comparaison des méthodes de désaisonnalisation mettait en jeu: (1) l'utilisation de la méthode de désaisonnalisation X-11-ARMMI avec et sans l'option d'extrapolation ARMMI; (2) l'emploi de facteurs saisonniers concourants ou prévus pour l'année à venir, pour la désaisonnalisation courante; et (3) le niveau d'agrégation auquel la désaisonnalisation doit être effectuée. Le choix de la ''meilleure'' procédure a été fait en fonction de deux propriétés extrêmement souhaitables, la minimisation des révisions et la maximisation de la régularité de la série désaisonnalisée.

Les résultats empiriques exposés au chapitre 5 indiquaient nettement que la méthode de désaisonnalisation à préférer pour la série de l'Indice des prix à la consommation de l'Ensemble se caractérisait ainsi:

(1) désaisonnalisation directe de la série agrégative (suivie, au second rang, par la désaisonnalisation indirecte portant sur les sept composantes principales);

(2) utilisation de l'extrapolation ARMMI de la méthode X-11-ARMMI;

(3) utilisation de facteurs saisonniers concourants pour la désaisonnalisation courante; et

(4) calcul des variations mensuelles d'après les premières estimations révisées du mois précédent.

Cependant, dans le choix de la désaisonnalisation directe de l'IPC de l'Ensemble, nous devrions toujours tenir compte de l'effet que la mise à jour de paniers fixes peut avoir sur la structure saisonnière.

Pendant la période analysée, aucun changement brutal appréciable n'a été observé, peut-être parce qu'il n'y avait alors que deux profils de dépenses observés. La fréquence des révisions de coefficients qui minimiserait les changements brusques de la composante saisonnière et l'utilisation d'une pondération variable plutôt que constante pour les articles saisonniers sont des sujets très importants qui méritent des recherches plus approfondies.

Une autre question pertinente qui n'a pas été traitée ici porte sur les méthodes de déclaration des données. La variabilité des taux de croissance des séries économiques dépend en partie de la fréquence de la communication des données (jour, semaine, mois, trimestre, etc.) et en partie de la durée de l'intervalle sur lequel le rythme de changement est mesuré. En général, la variabilité des taux diminue lorsque la durée sur laquelle le changement est mesuré passe à, disons, trois, six ou 12 mois.

On peut concevoir les méthodes de mesure du taux de croissance de manière à profiter de la réduction de la variance à la fois lorsque la fréquence de déclaration diminue et que l'intervalle entre les unités augmente.

Par ailleurs, les taux de croissance mesurés sur des périodes plus courtes ont l'avantage de signaler rapidement les points de retournement si les séries sont régulières. Cependant, si ces dernières présentent une forte variabilité, des taux mesurés sur de courtes périodes donnent beaucoup de faux signaux et il faut du temps pour discerner le vrai du faux. Une autre solution consiste à envisager des séries désaisonnalisées ''filtrées'', dans lesquelles le filtre a éliminé la plupart des variations irrégulières, ne laissant que les fluctuations de la tendance-cycle. Des recherches sont manifestement nécessaires sur ces questions importantes et elles devraient être fortement encouragées.

Renvoi

[1] Nous posons que le modèle de décomposition de la série X est multiplicatif, c'est-à-dire que $X = CSI$ lorsque C est la tendance-cycle, S le facteur saisonnier et I la fluctuation irrégulière.

Référence

Akaike, H. et Ishiguro, M., [1980], *BAYSEA, a Bayesian Seasonal Adjustment Program*, monographie de sciences informatiques n° 13, Institut de mathématique statistique, Tokyo, Japon.

Bayer, A. et Wilcox, D. [1981], *An Evaluation of Concurrent Seasonal Adjustment*, Technical Report, Special Studies Section, Washington, D.C., Board of Governors of the Federal Reserve System.

Bilongo, R. et Carbone, R. [1981], *Seasonal Adjustment and Adaptive Forecasting*, document de recherche, Faculté des Sciences de l'Administration, Université Laval, Québec, Canada.

Box, G.E.P. et Jenkins, G.M. [1970], *Time Series Analysis: Forecasting and Control*, San Francisco, Holden Day.

Burman, J.P. [1980], *''Seasonal Adjustment by Signal Extraction'', Journal of the Royal Statistical Society*, série A, 143, Partie 3, pp.321-327.

Cleveland, W.S., Grambsch, P. et Terpenning, I. [1982], *''A Comparison of Concurrent and Forecast Seasonal Adjustment''*, dans *Proceedings of the Business and Economics Section, Annual Meetings of the American Statistical Association* (à paraître).

Dagum, E.B. [1975], *Seasonal Factor Forecasts from ARIMA Models*, contributions libres de l'Institut statistique international, 40^e^ session vol. 3 pp.206-219.

__________ [1978], *Comparison and Assessment of Seasonal Adjustment Methods for Labour Force Series*, Washington, D.C., U.S. Government Printing Office.

__________ [1979], *The Seasonal Adjustment of Economic Time Series Aggregates. A Case Study, The Unemployment Rate*, Washington, D.C., U.S. Government Printing Office.

_________ [1980], *La méthode de désaisonnalisation X-11-ARMMI*, Ottawa, Statistique Canada, n° de catalogue 12-564F.

_________ [1982.a], *''Revision of Time Varying Seasonal Filters, Journal of Forecasting*, vol. 1, pp.173-187.

_________ [1982.b], *''The Effects of Asymmetric Filters on Seasonal Factor Revisions''*, *Journal of the American Statistical Association*, vol. 77, No. 380, pp.732-738.

_________ [1982.c], *''Revisions of Seasonally Adjusted Data Due to Filter Changes''*, dans *Proceedings of the Business and Economics Section, Annual Meetings of the American Statistical Association*, pp.39-45.

Duvall, R.M. [1966], *''Time Series Analysis by Modified Least-Squares Techniques'', Journal of the American Statistical Association*, vol. 61, pp.152-165.

Farley, D. et Zeller, S.H. [1978], *''Comments''*, dans *Seasonal Analysis of Economic Time Series*, (Arnold Zellner, ed.), Washington, D.C., U.S. Government Printing Office, pp.464-467.

Gersovitz, M. et MacKinnon, J.G. [1978], *''Seasonality in Regression: An Application of Smoothness Priors'', Journal of the American Statistical Association*, vol. 73, pp.264-273.

Geweke, J. [1978]: *''Temporal and Sectoral Aggregations of Seasonally Adjusted Time Series''*, dans *Seasonal Analysis of Economic Time Series* (Arnold Zellner, ed.) Washington, D.C., U.S. Government Printing Office, pp.411-427.

Granger, C.W.J. [1978], *''Seasonality: Causation, Interpretation and Implications''* dans *Seasonal Analysis of Economic Time Series* (Arnold Zellner, ed.) Washington, D.C., U.S. Government Printing Office, pp.33-46.

Hannan, E.J. [1960], *''The Estimation of Seasonal Variations'', The Australian Journal of Statistics*, vol. 2, pp.1-15.

Hatanaka, M. et Suzuki, M. [1967], *''A Theory of the Pseudo-spectrum and its Applications to Non-Stationary Dynamic Econometric Models''* dans *Essays in Mathematical Economics in Honor of Oskar Morgenstern* (Martin Shubik, ed.) New Jersey, Princeton University Press.

Havenner, A. et Swamy, P.A.V.B. [1978], *''A Random Coefficient Approach to Seasonal Adjustment of Economic Time Series''*, Special Studies Paper 124, Research and Statistics Division, Washington, D.C., Board of Governors of the Federal Reserve System.

Henshaw, R.C. [1966], *''Application of the General Linear Model to Seasonal Adjustment of Economic Time Series'', Econometrica*, vol. 34, pp.381-95.

Hillmer, S.C. et Tiao, G.C. [1982], *''An ARIMA-Model Based Approach to Seasonal Adjustment'', Journal of the American Statistical Association*, vol. 77, n° 377, pp.63-70.

Jorgenson, D.W. [1964], *''Minimum Variance Linear Unbiased Seasonal Adjustment of Economic Time Series'', Journal of the American Statistical Association*, vol. 59, pp.681-723.

Kenny, P. et Durbin, J. [1982], *''Local Trend Estimation and Seasonal Adjustment of Economic Time Series'', Journal of the Royal Statistical Society*, série A, 145, Partie 1, pp.1-41.

Kolmogorov, A.N. [1939], *''Sur l'interpolation des suites stationnaires''* Comptes rendus de l'Académie des sciences, Paris, vol. 208, p.2043.

Kuiper, J. [1978], *''A Survey and Comparative Analysis of Various Methods of Seasonal Adjustment''*, dans *Seasonal Analysis of Economic Time Series* (Arnold Zellner, ed.) Washington, D.C., U.S. Government Printing Office, pp.59-76.

__________ [1981], *''The Treatment of Extreme Value in the X-11-ARIMA Program''* dans *Time Series Analysis and Forecasting* (Anderson, O. et Perryman, M.R., ed.) Amsterdam, North-Holland Publishing Co., pp.257-66.

Ladd, G.W. [1964], *''Regression Analysis of Seasonal Data''*, *Journal of the American Statistical Association*, vol. 59, pp.402-21.

Lothian, J. et Morry, M. [1977], *''Le Problème de l'agrégation directe ou indirecte*, document de recherche, Division de la recherche et de l'analyse des chroniques, Ottawa, Statistique Canada.

Lovell, M.C. [1963], *''Seasonal Adjustment of Economic Time Series and Multiple Regression Analysis''*, *Journal of the American Statistical Association*, vol. 58, pp.963-1010.

Maravall, A. [1981], *''Errores de Medicion del Crecimiento a Corto Plazo de Series Monetarias Desestacionalizadas''*, Servicio de Estudios Economicos n° 21, Madrid, Banco de Espana.

McKenzie, S. [1982], *''An Evaluation of Concurrent Adjustment on Census Bureau Time Series''*, dans *Proceedings of the Business and Economics Section*, Annual Meetings of the American Statistical Association (à paraître).

Nerlove, M. [1964], *''Spectral Analysis of Seasonal Adjustment Procedures''*, *Econometrica*, vol. 32, pp.241-286.

________ Grether, D.M. et Carvalho, J.L. [1979], *Analysis of Economic Time Series*, New York, Academic Press.

Pierce, D. [1978], *''Seasonal Adjustment When Both Deterministic and Stochastic Seasonality are Present''*, dans *Seasonal Analysis of Economic* Time Series, (Arnold Zellner, ed.) Washington, D.C., U.S. Government Printing Office, pp.242-269.

________ [1980], *''Data Revision with Moving Average Seasonal Adjustment Procedures''*, *Journal of Econometrics*, vol. 14, n° 1, pp.95-114.

Rosenblatt, H.M. [1963], *''Spectral Analysis and Parametric Methods for Seasonal Adjustment of Economic Time Series''*, *Proceedings of the Business and Economics Section* of the Annual Meeting of the American Statistical Association, p.94.

Shiskin, J., Young, A.H. et Musgrave, J.C. [1967], *''The X-11 Variant of Census Method II Seasonal Adjustment Program''*, Technical Paper No. 15, Washington, D.C., U.S. Bureau of Census.

Statistique Canada [1982], *Document de référence de l'Indice des prix à la consommation, concepts et procédés*, n° de catalogue 62-553F.

Stephenson, J.A. et Farr, H.T. [1972], *''Seasonal Adjustment of Economic Data by Application of the General Linear Statistical Model'', Journal of the American Statistical Association*, vol. 69, pp.18-31.

Whittle, P. [1963], *Prediction and Regulation by Linear Least-Squares Methods*, London: English University Press.

Wiener, N. [1949], *Extrapolation, Interpolation and Smoothing of Stationary Time Series*, New York, John Wiley.

Wold, H. [1938], *A Study in the Analysis of Stationary Time Series*, Stockholm: Almquist and Wiksell (2e édition, 1954).

Young, A.H. [1968], *''Linear Approximations to the Census and BLS Seasonal Adjustment Methods'', Journal of the American Statistical Association*, vol. 63, pp.445-57

Indice des prix à la consommation de l'Emsemble

	Jan.	Fév	Mars	Avr.	Mai	Juin	Juil.	Août	Sept.	Oct.	Nov.	Déc.
1979	146.2	147.5	149.3	150.3	151.8	152.5	153.7	154.2	155.6	156.7	158.2	159.2
1980	160.1	161.4	163.2	164.2	166.1	167.9	169.2	170.8	172.3	173.8	176.0	177.0
1981	179.3	181.1	183.5	184.9	186.6	189.4	191.1	192.5	193.9	195.8	197.5	198.4

Indice des prix à la consommation de l'Ensemble, série désaisonnalisée par la méthode directe

	Jan.	Fév	Mars	Avr.	Mai	Juin	Juil.	Août	Sept.	Oct.	Nov.	Déc.
1979	146.7	147.9	149.2	150.5	151.6	152.1	153.3	154.0	155.6	156.8	158.1	159.6
1980	160.5	161.8	163.0	164.4	165.9	167.5	168.7	170.6	172.4	173.9	175.9	177.5
1981	179.8	181.5	183.3	185.1	186.4	188.9	190.6	192.3	194.0	195.9	197.3	199.0

Indice des prix à la consommation de l'Ensemble, série désaisonnalisée par la méthode indirecte

	Jan.	Fév	Mars	Avr.	Mai	Juin	Juil.	Août	Sept.	Oct.	Nov.	Déc.
1979	142.6	143.9	145.7	146.7	147.9	148.5	149.4	150.1	151.9	152.9	154.2	155.4
1980	156.2	151.5	159.2	160.3	161.8	163.5	164.5	166.2	166.2	169.7	171.6	172.9
1981	175.0	176.7	179.2	180.6	181.8	184.5	185.7	187.4	189.3	191.1	192.4	193.6

Indice des prix à la consommation de l'Ensemble, série désaisonnalisée par la méthode indirecte appliquée à sept composantes

	Jan.	Fév.	Mars	Avr.	Mai	Juin	Juil.	Août	Sept.	Oct.	Nov.	Déc.
1979	146.5	147.7	149.1	150.5	151.4	152.0	153.3	154.0	155.6	156.5	158.0	159.3
1980	160.4	161.6	163.0	164.2	165.8	167.5	168.7	170.6	172.2	173.8	175.6	177.3
1981	179.8	181.4	183.2	184.9	186.4	188.9	190.6	192.3	193.8	195.7	197.1	198.6

Indice des prix à la consommation – Aliments, série non corrigée

	Jan.	Fév.	Mars	Avr.	Mai	Juin	Juil.	Août	Sept.	Oct.	Nov.	Déc.
1979	153.6	157.3	161.3	162.8	163.7	164.5	167.4	166.2	166.5	167.6	168.1	170.4
1980	170.8	173.2	174.8	175.4	177.0	181.0	182.6	185.1	188.1	188.9	191.0	193.2
1981	194.2	197.5	198.8	200.8	199.9	203.5	206.2	206.8	206.3	206.1	205.7	204.0

Indice des prix à la consommation – Alimentation, série désaisonnalisée par la méthode directe

	Jan.	Fév.	Mars	Avr.	Mai	Juin	Juil.	Août	Sept.	Oct.	Nov.	Déc.
1979	155.0	157.9	161.8	163.3	163.2	163.2	165.0	164.7	166.2	168.2	169.4	171.5
1980	172.5	173.7	175.1	175.7	176.8	179.7	180.1	183.6	187.6	189.4	192.2	194.5
1981	196.2	198.0	199.1	201.1	200.0	202.2	203.6	205.2	205.7	206.5	206.7	205.4

Indice des prix à la consommation – Aliments, série désaisonnalisée par la méthode indirecte

	Jan.	Fév.	Mars	Avr.	Mai	Juin	Juil.	Août	Sept.	Oct.	Nov.	Déc.
1979	151.9	155.1	158.7	160.3	160.7	160.7	162.1	161.5	163.3	165.0	165.9	168.4
1980	169.0	170.8	172.1	172.6	173.8	176.8	176.8	179.7	184.8	186.1	188.7	191.1
1981	192.1	194.6	195.8	197.8	196.6	198.8	199.6	201.2	202.3	202.7	203.0	201.6

Indice des prix à la consommation – Ensemble sans les aliments, série non corrigée

	Jan.	Fév.	Mars	Avr.	Mai	Juin	Juil.	Aoûr	Sept.	Oct.	Nov.	Déc.
1979	143.4	144.2	145.3	146.1	147.7	148.4	149.2	150.3	151.9	153.0	158.8	155.3
1980	156.3	157.5	159.3	160.4	162.2	163.5	164.8	166.1	167.1	168.9	171.1	171.8
1981	174.4	175.8	178.5	179.7	182.0	184.7	186.1	187.7	189.6	192.1	194.4	195.9

Indice des prix à la consommation – Ensemble sans les aliments, série désaisonnalisée par la méthode directe

	Jan.	Fév.	Mars	Avr.	Mai	Juin	Juil.	Août	Sept.	Oct.	Nov.	Déc.
1979	143.5	144.4	145.1	146.2	147.6	148.4	149.4	150.5	152.1	152.8	154.2	155.3
1980	156.4	157.7	159.1	160.6	162.1	163.5	164.9	166.4	167.3	168.7	170.4	171.8
1981	174.5	176.1	178.3	179.9	181.8	184.7	186.2	188.0	189.9	192.0	193.6	195.0

Indice des prix à la consommation – Ensemble sans les aliments, série désaisonnalisée par la méthode directe

	Jan.	Fév.	Mars	Avr.	Mai	Juin	Juil.	Août	Sept.	Oct.	Nov.	Déc.
1979	140.1	140.8	142.1	143.0	144.3	145.1	145.9	147.0	148.8	149.7	151.0	151.8
1980	152.7	153.8	155.7	156.9	158.5	159.9	161.1	162.5	163.7	165.0	166.9	168.0
1981	170.3	171.8	174.6	175.8	177.8	180.6	181.9	183.6	185.8	187.9	189.5	191.4

Indice des prix à la consommation – Ensemble sans les aliments, série désaisonnalisée par la méthode indirecte appliquée par la méthode indirecte appliquée à six composantes

	Jan.	Fév.	Mars	Avr.	Mai	Juin	Juil.	Août	Sept.	Oct.	Nov.	Déc.
1979	143.6	144.4	145.1	146.4	147.7	148.5	149.6	150.6	152.1	152.8	154.3	155.3
1980	156.4	157.9	159.0	160.5	162.2	163.6	164.8	166.3	167.4	168.7	170.4	171.9
1981	174.6	176.2	178.3	179.8	181.9	184.6	186.4	188.1	189.8	192.2	193.7	196.1

Indice des prix à la consommation de l'Ensemble – Série désaisonnalisée avec extrapolation ARMMI et facteurs saisonniers concourants

	Jan.	Fév.	Mars	Avr.	Mai	Juin	Juil.	Août	Sept.	Oct.	Nov.	Déc.
1979	146.6	147.9	149.2	150.7	151.6	152.2	153.0	153.9	155.7	156.6	158.1	159.4
1980	160.7	161.9	163.0	164.3	165.6	167.5	168.5	170.4	172.2	173.8	175.8	177.1
1981	179.7	181.4	183.3	185.0	186.4	188.9	190.6	192.3	194.0	196.0	197.5	199.0

Indice des prix à la consommation de l'Ensemble – Série désaisonnalisée avec extrapolation ARMMI et facteurs saisonniers prévus

	Jan.	Fév.	Mars	Avr.	Mai	Juin	Juil.	Août	Sept.	Oct.	Nov.	Déc.
1979	146.7	148.0	149.3	150.8	151.6	152.2	153.0	153.9	155.8	156.5	158.0	159.3
1980	160.6	161.8	163.0	164.5	165.7	167.6	168.6	170.6	172.4	174.0	176.0	177.3
1981	179.9	181.6	183.4	185.3	186.3	189.0	190.7	192.3	193.9	195.9	197.1	198.5

Indice des prix à la consommation de l'Ensemble – Série désaisonnalisée sans extrapolation et avec facteurs saisonniers concourants

	Jan.	Fév.	Mars	Avr.	Mai	Juin	Juil.	Août	Sept.	Oct.	Nov.	Déc.
1979	146.5	147.5	149.4	150.9	151.8	152.4	153.2	153.9	155.6	156.5	158.2	159.5
1980	160.7	161.8	163.0	164.4	165.7	167.6	168.5	170.6	172.4	173.9	176.0	177.2
1981	179.9	181.6	183.4	185.0	186.2	188.9	190.6	192.3	193.9	196.1	197.4	198.9

Indice des prix à la consommation de l'Ensemble – Série désaisonnlaisée sans extrapolation et avec facteurs saisonniers prévus

	Jan.	Fév.	Mars	Avr.	Mai	Juin	Juil.	Août	Sept.	Oct.	Nov.	Déc.
1979	146.5	147.8	149.5	150.9	151.8	152.4	153.3	153.9	155.6	156.4	157.9	159.2
1980	160.6	161.8	163.1	164.5	165.7	167.6	168.6	170.6	172.3	173.8	176.0	177.3
1981	179.9	181.6	183.4	185.2	186.2	188.9	190.5	192.3	193.9	196.0	197.4	198.6

Indice des prix à la consommation de l'Ensemble – Série désaisonnalisée avec extrapolation ARMMI et facteurs saisonniers concourants premiers chiffres révisés

	Jan.	Fév.	Mars	Avr.	Mai	Juin	Juil.	Aoûr	Sept.	Oct.	Nov.	Déc.
1979	146.7	147.9	149.3	150.5	151.4	152.1	153.0	154.0	155.6	156.8	158.2	159.5
1980	160.6	161.8	163.0	164.2	165.7	167.3	168.7	170.5	172.2	173.9	175.7	177.4
1981	179.6	181.4	183.2	184.9	186.2	188.8	190.6	192.2	194.0	195.9	197.3	

THE TREATMENT OF SEASONALITY IN THE COST-OF-LIVING INDEX

*W.E. Diewert**
Department of Economics
University of British Columbia

SUMMARY

The paper addresses the problem of constructing monthly (and annual) Consumer Price or Cost-of-Living Indexes given that certain seasonal commodities are available only at certain times of the year.

The paper suggests the following solution: (i) that each seasonal commodity be treated as a separate commodity, (ii) compare the prices and quantities of the current year, where seasonal goods are regarded as separate commodities, with the corresponding prices and quantities pertaining to a base year, and (iii) use normal index number techniques in this framework. Thus in this framework, separate seasonal price indexes are never constructed and hence there is no need to seasonally adjust the indexes; in effect, they are already seasonally adjusted.

The paper also presents an economic justification for the suggested procedure.

The paper also shows how the recommended procedure for constructing annual indexes that appear every month may be approximated by an appropriately weighted moving average of monthly indexes. A monthly index compares the prices and quantities of a single month with the corresponding prices and quantities of the same month in a base year.

All of the procedures are illustrated using some artificial seasonal data invented by Ralph Turvey.

RÉSUMÉ

La communication traite du problème de la construction d'indices mensuels (et annuels) des prix à la consommation ou du coût de la vie compte tenu de ce que certains biens saisonniers ne sont offerts qu'à certaines époques de l'année.

Nous proposons la solution suivante: (i) traiter chaque bien saisonnier comme bien distinct, (ii) comparer les prix et les quantités de l'année en cours, lorsque les biens saisonniers sont considérés comme biens distincts, avec les prix et quantités correspondants d'une année de base, et (iii) utiliser les techniques normales des nombres-indices dans ce cadre. Ainsi, dans ce cadre, on ne construit jamais d'indices de prix saisonniers distincts, de sorte qu'il n'est plus nécessaire de désaisonnaliser les indices; effectivement, ils sont déjà désaisonnalisés.

La procédure proposée ci-dessus est illustrée au moyen de certaines données saisonnières artificielles inventées par Turvey. Il s'agit de données fictives sur les prix et les quantités pour 5 biens saisonniers (pommes, pêches, raisins, fraises et oranges) pour 4 mois par année, de sorte qu'il y a 5 x 4 x 12 = 240 observations en tout. À la section 3, la procédure proposée est mise en oeuvre avec les données de Turvey, selon le principe de la base fixe et celui de l'enchaînement pour quatre formules communes de nombres-indices: les indices de Laspeyres, de Paasche, Idéal de Fisher et Translog.

La procédure proposée produit des indices annuels pour chaque mois de l'année. À la section 4 de la communication, nous donnons une justification économique de la procédure. La justification exige: (i) un comportement de la part du consommateur qui cherche à maximiser l'utilité ou le bien-être, (ii) une certaine forme fonctionnelle spéciale pour la fonction d'utilité ou de préférence du consommateur (séparabilité additive dans les fonctions de sous-utilité saisonnière) et (iii) une parfaite hypothèse prévisionnelle à l'égard des futurs prix des biens saisonniers.

La solution proposée au problème des biens saisonniers comporte le traitement de chaque bien comme s'il s'agissait de 12 biens distincts et la comparaison des nombres-indices pour 12 mois consécutifs avec les données des mois correspondants d'une année de base

La procédure proposée a pour effet d'accroître considérablement la dimensionnalité de l'espace du bien, de sorte que les calculs requis peuvent être difficiles à réaliser avec des centaines de biens. La section 5 de la communication démontre comment on peut arriver à une approximation exacte de la prodécure recommandée pour la construction d'indices annuels qui apparaissent chaque mois à l'aide d'une moyenne mobile convenablement pondérée des indices mensuels. Un indice mensuel met en comparaison les prix et les quantités d'un mois donné avec les prix et les quantités correspondants du même mois d'une année de base.

La conclusion résume la communication et présente certaines observations sur le problème de prévision, c.-à-d. le problème qui consiste à utiliser, mettons, un taux mensuel courant d'inflation (qui compare les prix du mois courant avec les prix du mois correspondant d'une année de base) afin de prédire ce que sera le taux moyen annuel d'inflation. Nous soutenons que le problème de prévision n'est pas plus difficile (et probablement beaucoup plus simple) lorsqu'on traite les biens saisonniers selon les techniques proposées plutôt que selon la procédure classique.

1. Introduction

One of the most vexing problems that occurs in the construction of a consumer price index or a Konüs [1924] cost-of-living index is the treatment of seasonal commodities. These are commodities that are available only at certain times of the year. Given that our task is to construct a monthly or quarterly CPI or cost-of-living index, how should seasonal commodities be treated?

Turvey [1979] has nicely illustrated the problem. He invented fictitious price and quantity data for five seasonal commodities (apples, peaches, grapes, strawberries and oranges) for four years by month, so that there are $5 x 4 x 12 = 240$ observations in all. His data are tabled in Table 1. The reader will note that at certain times of the year, peaches and strawberries are unavailable, and thus the price and quantity entries for those seasons are left blank.

Turvey sent his "primary" data to statistical agencies throughout the world, asking them to use their normal techniques to construct monthly (and annual average) price indexes. About 20 countries replied, and Turvey tabled their estimated indexes for his artificial data.

Turvey summarizes the results as follows:

"It will be seen that the monthly indices display very large differences, e.g., a range of 129.12-169.5 in June, while the range of the simple annual means is much smaller. It will also be seen that the indices vary as to the peak month or months."

Turvey [1979; p.13].

Turvey [1979; p.13] notes that seasonal items in the current Canadian and Swiss consumer price indexes have a total weighting of about 7% and 10% respectively, so the problem of which treatment of seasonal commodities is appropriate is not an empirically unimportant question.

In Section 2 below, we review briefly some suggested solutions to the seasonality problem.

In Section 3, we propose a method for treating seasonal commodities. The method has the advantage that year-to-year inflation rates will be accurately determined. However, the economic foundations for the month-to- month comparisons are weaker than for the year-to-year comparisons.

In Section 4, we attempt to provide an economic justification for our method of comparing months.

In Section 5, an alternative (two-stage) procedure for constructing seasonal indexes is considered and compared with the method suggested in Section 3. The two methods yield similar numerical results.

Section 6 concludes.

		Months											
		1	2	3	4	5	6	7	8	9	10	11	12
							1970						
Apples	p	1.14	1.17	1.17	1.40	1.64	1.75	1.83	1.92	1.38	1.10	1.09	1.10
	x	3,086	3,765	4,363	4,842	4,439	5,323	4,165	3,224	4,025	5,784	6,949	3,924
Peaches	p	–	–	–	–	–	3.15	2.53	1.76	1.73	1.94	–	–
	x	–	–	–	–	–	91	498	6,504	4,923	865	–	–
Grapes	p	2.48	2.75	5.07	5.00	4.98	4.78	3.48	2.01	1.42	1.39	1.75	2.02
	x	82	35	98	26	75	82	1,490	2,937	2,826	1,290	338	
Strawberries	p	–	–	–	–	5.13	3.48	3.27	–	–	–	–	–
	x	–	–	–	–	700	2,709	1,970	–	–	–	–	
Oranges	p	1.3	1.25	1.21	1.22	1.28	1.33	1.45	1.54	1.57	1.61	1.59	1.31
	x	10,266	9,656	7,940	5,110	4,089	3,362	3,396	2,406	2,486	3,222	6,958	9,762
							1971						
Apples	p	1.25	1.36	1.38	1.57	1.77	1.86	1.94	2.02	1.55	1.34	1.33	1.30
	x	3,415	4,127	4,771	5,290	4,986	5,869	4,671	3,534	4,509	6,299	7,753	4,285
Peaches	p	–	–	–	–	–	3.77	2.85	1.98	1.80	1.95	–	–
	x	–	–	–	–	–	98	548	6,964	5,370	932	–	–
Grapes	p	2.80	3.32	5.48	5.67	5.44	5.30	3.93	2.33	1.66	1.64	2.10	2.35
	x	85	32	10	8	53	80	94	1,583	3,021	2,984	1,308	354
Strawberries	p	–	–	–	–	5.68	3.72	3.78	–	–	–	–	–
	x	–	–	–	–	806	3,166	2,153	–	–	–	–	–
Oranges	p	1.35	1.36	1.37	1.44	1.51	1.56	1.66	1.74	1.76	1.77	1.76	1.50
	x	10,888	10,314	8,797	5,590	4,377	3,681	3,748	2,649	2,726	3,477	3,548	10,727

TABLE 1. Prices and Quantities of Fruits Sold – Concluded

		Months											
		1	2	3	4	5	6	7	8	9	10	11	12
							1972						
Apples	p	1.43	1.53	1.59	1.73	1.89	1.98	2.07	2.12	1.73	1.56	1.56	1.49
	x	3,742	4,518	5,134	5,738	5,498	6,420	5,157	3,881	4,917	6,872	8,490	5,211
Peaches	p	–	–	–	–	–	4.69	3.32	2.29	1.90	1.97	–	–
	x	–	–	–	–	–	1.04	604	7,378	5,839	1,006	–	–
Grapes	p	3.20	4.03	6.06	6.59	6.01	5.94	4.61	2.79	1.94	1.95	2.46	2.92
	x	88	34	11	8	70	87	103	1,668	3,118	3,043	1,441	373
Strawberries	p	–	–	–	–	6.21	3.98	4.30	–	–	–	–	–
	x	–	–	–	–	931	3,642	2,533	–	–	–	–	–
Oranges	p	1.56	1.53	1.55	1.62	1.70	1.78	1.89	1.91	1.92	1.95	1.94	1.64
	x	11,569	10,993	9,621	6,063	4,625	3,970	4,078	2,883	2,957	3,759	8,238	11,827
							1973						
Apples	p	1.67	1.79	1.85	1.94	2.06	2.13	2.22	2.25	1.95	1.87	1.88	1.73
	x	4,051	4,909	5,567	6,253	6,101	7,023	5,671	4,187	5,446	7,377	9,283	4,955
Peaches	p	–	–	–	–	–	6.10	4.08	2.80	2.06	2.01	–	–
	x	–	–	–	–	–	111	653	7,856	6,291	1,073	–	–
Grapes	p	3.52	4.67	6.48	7.34	6.51	6.43	5.00	3.07	2.20	2.19	2.74	3.13
	x	91	37	11	9	80	92	97	1,754	3,208	3,199	1,646	391
Strawberries	p	–	–	–	–	6.89	4.32	4.91	–	–	–	–	–
	x	–	–	–	–	1,033	4,085	2,877	–	–	–	–	–
Oranges	p	1.68	1.66	1.70	1.85	1.95	2.03	2.10	2.12	2.07	2.13	2.14	1.79
	x	12,206	11,698	10,438	6,593	4,926	4,307	4,418	3,165	3,211	4,007	8,833	12,558

2. Suggested Solutions to the Seasonality Problem

We consider four approaches to the seasonal adjustment problems that have been sug-
gested in the literature.

The first approach is to drop seasonal items from the index altogether.[1] Needless to say, this is not a conceptually satisfying solution.

A second approach would be to use some sort of mechanical moving average or statistical method to smooth the series. Examples of this kind of technique may be found in Stone [1956; pp.77-88] and Allen [1975; pp.169-176]. This time series approach to seasonal adjustment procedures is very ably surveyed by Kuiper [1978].[2] The problem with this class of approaches is that it lacks an economic foundation.

A third approach is to invent shadow prices for seasonal commodities for the periods when they are out of season. The idea here is that the consumer has stable monthly preferences. The appropriate shadow price for some out-of-season commodities is the price that would make the consumer want to purchase 0 units of the commodity, given his income and other prices for that period. If there are not too many seasonal commodities, and if appropriate microeconomic data were available, then econometric techniques could be used to estimate these shadow prices.[3] However, the amount of effort required to implement this approach is probably too large. It is easier to make guesses about these shadow or reservation prices. Thus at present, Statistics Canada [1982; p.104] assumes that the out-of-season shadow price is equal to the last in-season price that is available. From the viewpoint of economic theory, under normal conditions, this estimated shadow price will be too low.

A fourth approach is to directly compare this year's season with the corresponding season in the previous year (or a base year).[4] This is fine as far as it goes, but it does not solve the problem of comparing one monthly index with the next monthly index.[5]

None of the above approaches seems very satisfactory from the viewpoint of economics. What then should we do?

3. A Proposed Solution to the Problem of Seasonality

The starting point for our solution to the seasonality problem is the following observation: there is no seasonality problem if we always compare the current year's prices and quantities with the corresponding prices and quantities of a base year. This amounts to treating seasonal goods as separate goods, so instead of thinking of monthly preferences that are defined over the five goods listed in Turvey's data, we think in terms of annual preferences defined over 5x12 = 60 goods.[6]

This way of looking at the problem of constructing annual indexes with seasonal data has been suggested by Stone [1956; p.75] and Diewert [1980; p.508] (and probably by many others).

Consider the data in Table 1. Arrange the 1970 price data into a vector of dimension 44 (we have dropped the 16 unavailable goods) and call it $p^1 \equiv (p_1^1,\ldots,p_{44}^1)$. Call the corresponding 1970 quantity vector $x^1 \equiv (x_1^1,\ldots,x_{44}^1)$. Define the 1971, 1972 and 1973 corresponding price and quantity vectors by (p^2,x^2), (p^3,x^3) and (p^4,x^4) respectively. It will also prove convenient to define base year price and quantity vectors p^0 and x^0. We shall define $p^0 \equiv p^1$ and $x^0 \equiv x^1$ so that our base year is really 1970.

Define a *price index formula* $P(p^0,p^t,x^0,x^t)$ to be some given function of the base period price and quantity vectors, p^0 and x^0, and of the current period price and quantity vectors, p^t and x^t.

In Diewert [1983], we found that four index number formulae could provide approximations to an underlying cost-of-living index. These four indexes were the Laspeyres index P_L, the Paasche index P_P, the Fisher ideal index P_2, and the translog index P_T. These functions are defined below.[7]

$$P_L(p^0,p^t,x^0,x^t) \equiv p^t \cdot x^0 / p^0 \cdot x^0, \tag{1}$$

$$P_P(p^0,p^t,x^0,x^t) \equiv p^t \cdot x^t / p^0 \cdot x^t, \tag{2}$$

$$P_F(p^0,p^t,x^0,x^t) \equiv [(p^t \cdot x^0 / p^0 \cdot x^0)(p^t \cdot x^t / p^0 \cdot x^t)]^{1/2}, \text{ and} \tag{3}$$

$$P_T(p^0,p^t,x^0,x^t) \equiv \prod_{i=1}^{44} (p_i^t / p_i^0)^{[1/2][s_i^0 + s_i^t]} \tag{4}$$

where $s_i^0 \equiv p_i^0 x_i^0 / p^0 \cdot x^0$ and $s_i^t \equiv p_i^t x_i^t / p^t \cdot x^t$ for $i = 1,\ldots,44$.

Using the formulae (1)-(4) and the data in Table 1 we table the resulting annual index numbers for t = 1, 2, 3, 4.

TABLE 2. Annual Fixed Base Price Levels Using Alternative Index Number Formulae

Year	t	P_L Laspeyres	P_P Paasche	P_F Fisher	P_T Translog
1970	1	1.00000	1.00000	1.00000	1.00000
1971	2	1.11945	1.11956	1.11950	1.11952
1972	3	1.25263	1.25238	1.25200	1.25200
1973	4	1.40296	1.40050	1.40173	1.40175

Table 2 indicates that the annual indexes are remarkably consistent across functional forms. As we would expect from theoretical considerations, P_F and P_T are very close to each other and they are both bracketed by P_L and P_P.

There may be some slight rounding errors in our computations (although a double precision program was used) since Turvey [1979; p.13] computed the Laspeyres fixed base index for 1973 to be 1.4028 compared to our 1.40296.

It is also possible to compute the annual indexes using the chain principle. Using this principle, we construct the chain index for the formula P for year 3 say as the product of the year-by-year links:

$$P(p^0,p^1,x^0,x^1)P(p^1,p^2,x^1,x^2)P(p^2,p^3,x^2,x^3). \tag{5}$$

Using the chain principle and the data in Table 1, the resulting annual index numbers are tabled in Table 3.

TABLE 3. Annual Chained Price Levels Using Alternative Index Number Formulae

Year	t	P_L Laspeyres	P_P Paasche	P_F Fisher	P_T Translog
1970	1	1.00000	1.00000	1.00000	1.00000
1971	2	1.11945	1.11956	1.11950	1.11952
1972	3	1.25253	1.25174	1.25214	1.25214
1973	4	1.40259	1.40154	1.40207	1.40207

Note that Row 1 in Tables 2 and 3 is a row of ones. This follows from our addition of year 0 to the data, since we took the data for year 0 to equal the corresponding data for year 1. Also note that Row 2 in Tables 2 and 3 are equal since the chain principle has not yet been able to generate different numbers. Comparing the last two rows of Tables 2 and 3 leads us to conclude that the Turvey data generates essentially the same annual indexes for both the fixed base and chained indexes, irrespective of the four functional forms that we have chosen to work with.[8]

To make further progress, we next observe that there is no reason why we should always compare the base year's January-to-December with the current year's January-to-December observations. Why not compare February 1970 through December 1970 plus January 1971 with February 1971 through December 1971 plus January 1971? If we continue on in this manner, the resulting annual index number series will be just as consistent with economic theory as the usual calendar year comparisons that we have just tabled.

Thus we can now conceive of 12 different annual index number comparisons. It is convenient to define some new notation that will enable us to describe with more precision what we are suggesting. Look back at Table 1, and add the year 1969 to the top of the table, and use the 1970 entries as 1969 entries. Now we have price and quantity vectors (of varying dimension depending on how many goods are actually available during that month) for 60 months. Label these 60 price vectors as $p^1, p^2, \ldots p^{60}$, where p^1 represents the three January 1969 prices, p^2 represents the three February 1969 prices,..., p^6 represents the five June 1969 prices,..., p^{13} represents the three January 1970 prices, p^{14} represents the three February 1970 prices,..., and p^{60} represents the three December 1973 prices. Define the corresponding 60 monthly quantity vectors as $x^1, x^2, \ldots, x^{60}$. Let P be one of the index number formulae defined by (1)-(4). The comparisons in Table 2 are of the form as below:

$$P(p^1,p^2,\ldots,p^{12},p^{13},\ldots,p^{24};\ x^1,x^2,\ldots,x^{12};\ x^{13},x^{14},\ldots,x^{24}) \tag{6}$$

$$P(p^1,p^2,\ldots,p^{12};p^{25},\ldots,p^{36};\ x^1,x^2,\ldots,x^{12};x^{25},x^{26},\ldots,x^{36})$$

etc.

Now we suggest making comparisons of the form:

$$P(p^2,p^3,\ldots,p^{13};\ p^{14},p^{15},\ldots,p^{25};\ x^2,x^3,\ldots,x^{13};\ x^{14},x^{15},\ldots,x^{25}) \tag{7}$$

$$P(p^2,p^3,\ldots,p^{13};p^{26},p^{27},\ldots,p^{37};\ x^2,x^3,\ldots,x^{13};\ x^{26},x^{27},\ldots,x^{37})$$

etc.

The annual comparisons that we make using the split year base in (7) are just as valid as the normal annual comparisons that we made using (6), (or at least, they would be just as valid if we had not introduced our "dummy" year of observations). This is well and fine, but we have still not solved the problem of making valid monthly comparisons between these annual (split year base) indexes.

This is where the "dummy" year comes into its own. If we look at the first index number in (7), it is comparing the level of prices in months two to 13 with months 14 to 25. If we recall that months one to 12 are identical to months 13 to 24, it can be seen that what

we are really doing is comparing the 1970 prices and quantities (for all 12 months) wit the prices and quantities of February through December of 1970 **plus** the prices and quan tities of January 1971; this effectively isolates the contribution (to the inflation rate) o adding January 1971 to our initial 1970 base months (and dropping January 1970 of course) Thus the entry in the first line of (7) provides us with our desired link between the January to-January annual indexes and the February-to-February annual indexes.

Thus our complete set of fixed base index number computations for a representativ price index formula P will run as follows:

$$P(p^1,\ldots,p^{12};\ p^{13},\ldots,p^{24};\ x^1,\ldots,x^{12};\ x^{13},\ldots,x^{24}) = 1 \tag{8}$$

$$P(p^1,\ldots,p^{12};\ p^{25},p^{14},p^{15},\ldots,p^{24};\ x^1,\ldots,x^{12};\ x^{25},x^{14},\ldots,x^{24})$$

$$P(p^1,\ldots,p^{12};\ p^{25},p^{26},p^{15},\ldots,p^{24};\ x^1,\ldots,x^{12};\ x^{25},x^{26},\ldots,x^{24})$$

$$P(p^1,\ldots,p^{12};\ p^{49},p^{50},\ldots,p^{60};\ x^1,\ldots,x^{12};\ x^{49},x^{50},\ldots,x^{60}).$$

Thus each run of 12 consecutive months is compared with the corresponding month in the base year. The resulting index number computations for our four formulae are tabl ed in Table 4.

The reader will note that the annual inflation rates of Table 2 appear in Table 4 (se 1970:12, 1971:12, 1972:12 and 1973:12). The entries for the year 1971 in Table 4 are ir a sense artificial as we explained above, but the remaining entries have a solid economi interpretation as the level of prices for a 12-month period relative to the level of price for the base 12 months. The first entries may also be interpreted in this way, but the microeconomic foundations of this interpretation are not as firm, since for the first 12 entries, the base 12 months overlap with the current 12 months. We shall return to the problem of providing a possible microeconomic interpretation for these anomalous entries in the following section.

Once we have constructed the first 12 entries in Table 4, we may switch to the chain principle rather than maintaining the fixed base throughout. The resulting chained index

umbers for our four formulae are tabled in Table 5. The reader will note that the first 2 entries in Table 5 are the same as the corresponding entries in Table 4. Moreover, the ntries 1970:12, 1971:12, 1972:12 and 1973:12 are the same as the entries in Table 3.

ABLE 4. Fixed Base Year-to-Year Price Levels by Month

ear	Month	Laspeyres	Paasche	Fisher	Translog
970	12	1.00000	1.00000	1.00000	1.00000
971	1	1.00391	1.00420	1.00405	1.00406
	2	1.01192	1.01272	1.01232	1.01234
	3	1.02168	1.02320	1.02244	1.02249
	4	1.03039	1.03238	1.03138	1.03143
	5	1.03892	1.04127	1.04010	1.04014
	6	1.04830	1.05078	1.04954	1.04957
	7	1.05888	1.06136	1.06012	1.06013
	8	1.07096	1.07310	1.07203	1.07202
	9	1.08080	1.08242	1.08161	1.08159
	10	1.09247	1.09364	1.09305	1.09303
	11	1.10719	1.10773	1.10746	1.10746
	12	1.11945	1.11956	1.11950	1.11952
972	1	1.13164	1.13228	1.13196	1.13198
	2	1.14193	1.14319	1.14256	1.14259
	3	1.15240	1.15479	1.15360	1.15366
	4	1.15999	1.16302	1.16150	1.16157
	5	1.16755	1.17100	1.16927	1.16931
	6	1.17742	1.18090	1.17916	1.17917
	7	1.18917	1.19326	1.19121	1.19121
	8	1.20447	1.20806	1.20627	0.20624
	9	1.21534	1.21830	1.21677	1.21673
	10	1.22757	1.23011	1.22884	1.22879
	11	1.24235	1.24148	1.24191	1.24192
	12	1.25263	1.25138	1.25200	1.25200
973	1	1.26151	1.26039	1.26095	1.26095
	2	1.27157	1.27077	1.27117	1.27119
	3	1.28195	1.28206	1.28201	1.28206
	4	1.29175	1.29242	1.29208	1.29213
	5	1.30185	1.30272	1.30229	1.30230
	6	1.31400	1.31462	1.31431	1.31430
	7	1.32715	1.32845	1.32780	1.32780
	8	1.34791	1.34794	1.34793	1.34790
	9	1.36044	1.35919	1.35981	1.35979
	10	1.37419	1.37230	1.37325	1.37324
	11	1.39192	1.39002	1.39097	1.39100
	12	1.40296	1.40050	1.40173	1.40175

TABLE 5. Chained Year-to-Year Price Levels by Month

Year	Month	Laspeyres	Paasche	Fisher	Translog
1970	12	1.00000	1.00000	1.00000	1.00000
1971	1	1.00391	1.00420	1.00405	1.00406
	2	1.01192	1.01272	1.01232	1.01234
	3	1.02168	1.02320	1.02244	1.02249
	4	1.03039	1.03238	1.03138	1.03143
	5	1.03982	1.04127	1.04010	1.04014
	6	1.04830	1.05078	1.04954	1.04957
	7	1.05888	1.06136	1.06012	1.06013
	8	1.07096	1.07310	1.07203	1.07202
	9	1.08080	1.08242	1.08161	1.08159
	10	1.09247	1.09364	1.09305	1.09303
	11	1.10719	1.10773	1.10746	1.10746
	12	1.11945	1.11956	1.11950	1.11952
1972	1	1.13179	1.13224	1.13201	1.13203
	2	1.14204	1.14304	1.14254	1.14258
	3	1.15260	1.15439	1.15349	1.15355
	4	1.16015	1.16242	1.16128	1.16135
	5	1.16743	1.17008	1.16875	1.16882
	6	1.17674	1.17951	1.17812	1.17816
	7	1.18839	1.19123	1.18981	1.18982
	8	1.20355	1.20594	1.20474	1.20474
	9	1.21414	1.21596	1.21505	1.21503
	10	1.22653	1.22776	1.22714	1.22712
	11	1.24215	1.24172	1.24193	1.24193
	12	1.25253	1.25174	1.25214	1.25214
1973	1	1.26155	1.26141	1.26148	1.26149
	2	1.27163	1.27220	1.27191	1.27194
	3	1.28210	1.28359	1.28284	1.28290
	4	1.29198	1.29397	1.29297	1.29304
	5	1.30171	1.30409	1.30290	1.30296
	6	1.31292	1.31550	1.31421	1.31425
	7	1.32601	1.32867	1.32734	1.32735
	8	1.34664	1.34872	1.34768	1.34766
	9	1.35873	1.36022	1.35947	1.35945
	10	1.37286	1.37373	1.37329	1.37326
	11	1.39136	1.39084	1.39110	1.39109
	12	1.40259	1.40154	1.40207	1.40207

A comparison of Tables 4 and 5 yields the same result that our comparison of Tables 2 and 3 yielded: for the Turvey data, it does not matter whether we use a fixed base principle or the chain principle.

Note that the data in Tables 4 and 5 do not exhibit wild monthly fluctuations. Hence, it seems that our suggested method does in fact remove the seasonal fluctuations in the Turvey data set.[9]

We turn now to the problem of justifying the first 12 entries in Tables 4 and 5.

4. An Economic Justification for the First Year's Entries

Consider the Paasche and Laspeyres index number formulae P_L and P_P defined by (1) and (2) over a pair of price and quantity vectors, (p^0,x^0) and (p^t,x^t). In order to use these index number formulae in order to place bounds on a consumer's true (Konüs [1924]) cost-of-living index,[10] we need to assume that the consumer has well-defined preferences over the goods (summarized by the utility function $u = F(x)$ say) and that he is engaging in utility maximizing behaviour during the two periods in question. Specifically, we need to assume that x^0 is a solution to the utility maximization problem

$$\max_x [F(x): \ p^0 \cdot x \leq p^0 \cdot x^0] \tag{9}$$

and x^t is a solution to the utility maximization problem

$$\max_x \ [F(x): p^t \cdot x \leq p^t \cdot x^t]. \tag{10}$$

Now consider our seasonal problem. In order to simplify the notation, let us assume that a year has only two seasons. Following the notation used in the previous section, let x^1 be a vector of consumer purchases associated with the first season, and x^2 be a vector of consumption variables (of different dimension in general) associated with the second season. The consumer's annual utility function is assumed to be $F(x^1,x^2)$. Suppose that we can observe the consumer's purchases during year 1 and the first half of year 2; i.e.,

we are given the observed quantity vectors x^{1*}, x^{2*} and x^{3*} (where x^{1*} and x^{3*} are vectors of the same dimension). Suppose too that we can observe the corresponding seasonal price vectors p^1,p^2 and p^3. We may now compare the price level in the second half of year 1 and the first half of year 2 by evaluating one of our four index number formulae of the form:

$$P[(p^1,p^2);\ (p^3,p^2);\ (x^{1*},x^{2*}),\ (x^{3*},x^{2*})]. \tag{11}$$

In order to relate (11) to the consumer's true cost-of-living (i.e., in order to apply the traditional economic theory of index numbers), it would be necessary to assume that (x^{1*},x^{2*}) is a solution to the following year 1 utility maximization problem,

$$\max_{x^1,x^2}\left\{F(x^1,x^2):\ p^1\cdot x^1\cdot x^1+p^2\cdot x^2 \le p^1\cdot x^{1*}+p^2\cdot x^{2*}\right\} \tag{12}$$

and that x^{2*}, x^{3*} is a solution to the following split year utility maximization problem:

$$\max_{x^2,x^3}\left\{F(x^3,x^2):\ p^2\cdot x^2+p^3\cdot x^3 \le p^2\cdot x^{2*}+p^3\cdot x^{3*}\right\}. \tag{13}$$

Can we find conditions that are sufficient for (12) and (13) to hold? The answer is of course yes.

Let r_1 be the rate of interest between the first season and the second season, let r_2 be the rate of interest between the second and third seasons, and define the positive discount terms $\delta_1 \equiv 1/(1+r_1)$ and $\delta_2 \equiv 1/(1+r_2)$. Suppose x^{1*},x^{2*},x^{3*} is a solution to the following intertemporal utility maximization problem defined over the three seasons:

$$\max_{x^1,x^2,x^3}\left\{f^1(x^1)+f^2(x^2)+f^1(x^3):\ p^1\cdot x^1 + \delta_1 p^2\cdot x^2+\delta_1\delta_2 p^3\cdot x^3 \le p^1\cdot x^{1*}+\delta_1 p^2\cdot x^{2*}+\delta_1\delta_2 p^3\cdot x^{3*}\right\}. \tag{14}$$

Note that the consumer's intertemporal utility function is $f^1(x^1)+f^2(x^2)+f^1(x^3)$ and f^1 may be interpreted as a season 1 subutility function while f^2 is the season 2 subutility function. These two functions can be completely different.

Note also that the consumer must make his purchases of x^1 in the first season before the spot prices of commodities in seasons 2 and 3, p^2 and p^3, actually emerge. Hence, assuming that the consumer's observed *ex–post* purchases x^{1*},x^{2*},x^{3*} solve (14) means that the consumer is accurately forecasting these future spot prices when he makes his consumption purchases in season 1. We must acknowledge that this is a somewhat dubious assumption.

In any case, given our assumption that x^{1*},x^{2*},x^{3*} solves (14), we may set x^3 in (14) equal to x^{3*} and conclude

$$(x^{1*},x^{2*}) \text{ solves } \max_{x^1,x^2} \left\{f^1(x^1) + f^2(x^2)\colon p^1{\cdot}x^1 + \delta_1 p^2{\cdot}x^2 \le p^1{\cdot}x^{1*} + \delta_1 p^2{\cdot}x^{2*}\right\}. \tag{15}$$

We may also set $x^1 = x^{1*}$ in (14) and conclude that [11]

$$(x^{3*},x^{2*}) \text{ solves } \max_{x^2,x^3} \left\{f^1(x^3) + f^2(x^2)\colon \delta_2 p^3{\cdot}x^3 + p^2{\cdot}x^2 \le \delta_2 p^3{\cdot}x^{3*} + p^2{\cdot}x^{2*}\right\}. \tag{16}$$

It can be seen that (15) is almost (12) and (16) is almost (13). In order to obtain an exact equivalence, it is necessary to assume that $\delta_1 = \delta_2 = 1$; i.e., that the seasonal interest rates $r_1 = r_2 = 0$. Under this additional assumption, it can be seen that we may interpret index number formulae of the type (11) in the usual way.

Thus we have provided an economic justification for our "dummy" year procedure explained in the previous section. The reader can see that it is not a terribly strong one, but it is not negligible either.

It would be of interest to establish (12) and (13) for more general functional forms for F (other than $F(x^1,x^2) = f^1(x^1) + f^2(x^2)$ which was the functional form that we used). I leave this problem to the functional equation experts at this conference.

5. Two-Stage Aggregation

Our suggested solution to the problem of seasonal commodities involves treating each commodity as 12 commodities and then calculating index number formulae of the type (6). This has the effect of greatly increasing the dimensionality of the commodity space, and thus it may be numerically difficult to carry out the required computations when we are dealing with thousands of goods. However, it is possible to simplify the computations by using the two-stage aggregation theorem proved in Diewert [1978; pp.889-890]: by using a superlative index number formula such as P_F or P_T, a close approximation to the correct annual aggregate price defined by (3) or (4) above can be constructed by constructing monthly indexes and then aggregating over the year. Let us see how this theoretical result fares when it is subjected to the Turvey data. At the same time, we shall see how the two-stage aggregation procedure enables us to avoid doing the full set of computations that are inherent in the one-stage procedure of the form (6).

The first thing we do is to treat the observations in a given month separately. Thus we construct 12 sets of monthly index number comparisons, using our four functional forms for the index number formulae (Laspeyres, Paasche, Fisher and Translog). Thus we compare the January 1970 data with the January 1971, January 1972 and January 1973 data respectively. Then do the same comparisons for the February data, and so on. Thus we are constructing fixed base monthly indexes that compare the same month across years. The resulting comparisons may be found in Table 6. Note that we have set the indexes equal to 1 for each month in the base year 1970. The Laspeyres entry for 1971:1 is the Laspeyres price index for the January 1971 data in Table 1 compared with the January 1970 data, while the Laspeyres entry for 1972:1 compares the data for January 1972 with the data for January 1970. The Laspeyres entry for 1971:2 compares the February data for 1971 with the February data for 1970, and so on. Comparing the data in Table 6 with the data in Table 4, we note that the data in Table 4 is much smoother and much more comparable across formulae. This is to be expected since the index numbers in Table 4 are essentially averages over 12 months of data whereas the index numbers in Table 6 compare only one month with the same month in the base year.

TABLE 6. Fixed Base Monthly Price Indexes Across Years

Year	Month	Laspeyres	Paasche	Fisher	Translog
1970	1	1.00000	1.00000	1.00000	1.00000
	2	1.00000	1.00000	1.00000	1.00000
	3	1.00000	1.00000	1.00000	1.00000
	4	1.00000	1.00000	1.00000	1.00000
	5	1.00000	1.00000	1.00000	1.00000
	6	1.00000	1.00000	1.00000	1.00000
	7	1.00000	1.00000	1.00000	1.00000
	8	1.00000	1.00000	1.00000	1.00000
	9	1.00000	1.00000	1.00000	1.00000
	10	1.00000	1.00000	1.00000	1.00000
	11	1.00000	1.00000	1.00000	1.00000
	12	1.00000	1.00000	1.00000	1.00000
1971	1	1.05154	1.05188	1.05171	1.05172
	2	1.10847	1.10876	1.10861	1.10862
	3	1.14842	1.14828	1.14835	1.14835
	4	1.14960	1.14962	1.14961	1.14961
	5	1.11793	1.11676	1.11735	1.11734
	6	1.08820	1.08761	1.08790	1.08790
	7	1.11553	1.11508	1.11531	1.11531
	8	1.11140	1.11108	1.11124	1.11124
	9	1.09962	1.09905	1.09934	1.09933
	10	1.15265	1.15276	1.15270	1.15269
	11	1.15803	1.17836	1.16815	1.16808
	12	1.15467	1.15461	1.15464	1.15464
1972	1	1.21232	1.21289	1.21261	1.21261
	2	1.24765	1.24832	1.24798	1.24799
	3	1.30770	1.30718	1.30744	1.30744
	4	1.27998	1.27999	1.27998	1.27999
	5	1.22236	1.21906	1.22071	1.22066
	6	1.18101	1.17857	1.17979	1.17977
	7	1.24395	1.24430	1.24412	1.24410
	8	1.25247	1.24989	1.25118	1.25117
	9	1.20967	1.20688	1.20827	1.20821
	10	1.31270	1.31185	1.31227	1.31215
	11	1.31669	1.31746	1.31708	1.31709
	12	1.28425	1.28531	1.28478	1.28478
1973	1	1.32951	1.33218	1.33085	1.33091
	2	1.38382	1.38650	1.38516	1.38522
	3	1.46553	1.46438	1.46495	1.46495
	4	1.44838	1.44834	1.44836	1.44836
	5	1.36194	1.35456	1.35824	1.35813
	6	1.29523	1.29053	1.29288	1.29283
	7	1.38762	1.38895	1.38829	1.38823
	8	1.44389	1.43769	1.44079	1.44071
	9	1.33654	1.33177	1.33415	1.33401
	10	1.49256	1.49294	1.49275	1.49243
	11	1.50699	1.51108	1.50903	1.50909
	12	1.42351	1.42234	1.42292	1.42294

Statistical agencies often compute an annual CPI by taking a simple average (with weights 1/12) of 12 consecutive monthly indexes of the Laspeyres type that we have calculated in Table 6. In order to compare our recommended annual index numbers that are tabled in Tables 4 and 5 above with the "official" practice, we have computed the simple 12-month moving averages of the indexes listed in Table 6. (Actually, "official" monthly price indexes often use a fixed basket that is constant across months; hence perhaps we should call the resulting fixed weight moving average indexes "simple"). The results are displayed in Table 7. It is interesting to note that the data in Table 7 differ with the data in Table 4 by at most one per cent. Hence, "simple" annual indexes rather closely approximate our theoretically desirable indexes tabled in Table 4.

Return to the monthly data in Table 6. Add an additional 12-base year [1969] observations to the table, where each added observation is a one. Consider the resulting 60 numbers in any of the last four columns, say for definiteness, the Laspeyres column. Define corresponding aggregate commodities by

$$X_t \equiv p^t \cdot x^t / P_t;\ t = 1, \ldots, 60 \tag{17}$$

where the p^t and x^t are the monthly price and quantity vectors defined earlier (6). Now use the resulting aggregate price and quantities P_t and X_t in order to construct split year Laspeyres index number comparisons with the aggregate data of the base year; i.e., use the formula (8) with $P = P_L$, and the P_t replacing the p^t and the X_t replacing the x^t. The resulting 37 numbers are tabled in the Laspeyres column of Table 8. Repeat the above procedure using the data in the Paasche, Fisher and Translog columns of Table 6 in order to construct the corresponding X_t via (17). Then use (8) with $P = P_P$, P_F and P_T respectively and the corresponding P_t, X_t replacing the p^t, x^t. The resulting two stage Paasche, Fisher and Translog indexes are tabled in the last 3 columns of Table 8.

Note that the two-stage Fisher and Translog indexes in Table 8 coincide with the corresponding single-stage Fisher and Translog indexes in both Tables 4 and 5 to four significant figures. Hence, the two-stage indexes may be used in order to approximate very closely the corresponding single- stage indexes.

TABLE 7. 12-Month Moving Averages of the Fixed Base Monthly Indexes

Year	Month	Laspeyres	Paasche	Fisher	Translog
1970	12	1.00000	1.00000	1.00000	1.00000
1971	1	1.00430	1.00432	1.00431	1.00431
	2	1.01333	1.01339	1.01336	1.01336
	3	1.02570	1.02574	1.02572	1.02572
	4	1.03817	1.03821	1.03819	1.03819
	5	1.04800	1.04794	1.04797	1.04797
	6	1.05535	1.05524	1.05529	1.05529
	7	1.06497	1.06483	1.06490	1.06490
	8	1.07426	1.07409	1.07417	1.07417
	9	1.08256	1.08234	1.08245	1.08245
	10	1.09528	1.09507	1.09518	1.09517
	11	1.10845	1.10994	1.10919	1.10918
	12	1.12134	1.12282	1.12208	1.12207
1972	1	1.13474	1.13624	1.13548	1.13548
	2	1.14634	1.14787	1.14710	1.14709
	3	1.15961	1.16111	1.16036	1.16035
	4	1.17047	1.17197	1.17122	1.17121
	5	1.17918	1.18050	1.17983	1.17982
	6	1.18691	1.18808	1.18749	1.18748
	7	1.19761	1.19885	1.19823	1.19821
	8	1.20937	1.21042	1.20989	1.20987
	9	1.21854	1.21940	1.21897	1.21895
	10	1.23188	1.23266	1.23226	1.23223
	11	1.24510	1.24425	1.24467	1.24465
	12	1.25589	1.25514	1.25552	1.25550
1973	1	1.26566	1.26508	1.26537	1.26535
	2	1.27701	1.27660	1.27680	1.27679
	3	1.29016	1.28970	1.28993	1.28992
	4	1.30419	1.30373	1.30396	1.30395
	5	1.31583	1.31502	1.31542	1.31540
	6	1.32534	1.32435	1.32484	1.32483
	7	1.33732	1.33640	1.33686	1.33684
	8	1.35327	1.35205	1.35266	1.35263
	9	1.36384	1.36246	1.36315	1.36311
	10	1.37883	1.37755	1.37819	1.37814
	11	1.39469	1.39369	1.39419	1.39414
	12	1.40629	1.40510	1.40570	1.40565

TABLE 8. Two-Stage Fixed Base Year-to-Year Price Levels by Month

Year	Month	Laspeyres	Paasche	Fisher	Translog
1970	12	1.00000	1.00000	1.00000	1.00000
1971	1	1.00391	1.00375	1.00383	1.00383
	2	1.01192	1.01111	1.01152	1.01151
	3	1.02168	1.01987	1.02078	1.02075
	4	1.03039	1.02781	1.02910	1.02907
	5	1.03892	1.03586	1.03739	1.03736
	6	1.04830	1.04513	1.04671	1.04670
	7	1.05888	1.05556	1.05721	1.05721
	8	1.07096	1.06778	1.06937	1.06937
	9	1.08080	1.07802	1.07940	1.07941
	10	1.09247	1.08993	1.09119	1.09121
	11	1.10719	1.10694	1.10711	1.10711
	12	1.11945	1.12010	1.11981	1.11981
1972	1	1.13164	1.13297	1.13234	1.13234
	2	1.14193	1.14385	1.14292	1.14294
	3	1.15240	1.15512	1.15379	1.15383
	4	1.15999	1.16328	1.16167	1.16171
	5	1.16755	1.17108	1.16935	1.16937
	6	1.17742	1.18074	1.17912	1.17912
	7	1.18917	1.19263	1.19094	1.19093
	8	1.20447	1.20736	1.20596	1.20593
	9	1.21534	1.21751	1.21647	1.21644
	10	1.22757	1.22910	1.22838	1.22834
	11	1.24235	1.23944	1.24092	1.24087
	12	1.25263	1.24929	1.25098	1.25095
1973	1	1.26151	1.25877	1.26016	1.26015
	2	1.27157	1.26963	1.27063	1.27064
	3	1.28195	1.28123	1.28161	1.28168
	4	1.29175	1.29168	1.29174	1.29182
	5	1.30185	1.30227	1.30209	1.30215
	6	1.31400	1.31454	1.31431	1.31433
	7	1.32715	1.32837	1.32780	1.32779
	8	1.34791	1.34822	1.34810	1.34806
	9	1.36044	1.36006	1.36028	1.36024
	10	1.37419	1.37349	1.37387	1.37382
	11	1.39192	1.38952	1.39075	1.39069
	12	1.40296	1.39985	1.40143	1.40138

We have not tabled the analogous computations for the chained indexes; the results are very similar to the fixed base results.

6. Conclusion

Our suggested solution to the problem of constructing seasonal indexes may be summarized as follows:

(1) try to avoid single month-to-month comparisons; instead focus on annual indexes where the prices and quantities of any consecutive run of seasons that make up a year is compared with the corresponding seasonal prices and quantities in a base year;

(2) the above method gives valid year-to-year comparisons for all runs of seasons except the initial runs where we are comparing combinations of base year seasons plus next year seasons with the entire base year (the split year comparison problem);

(3) under additional assumptions, the split year comparisons may also be justified from an economic point of view.

When our suggested solution is implemented on the Turvey data, we find that the Laspeyres, Paasche, Fisher and Translog formulae all give the same answer to three significant figures, and this is true whether we use a fixed base or the chain principle or whether we use a properly formulated two-stage procedure (see Tables 4, 5 and 8). Thus the results of Généreux [1983] and Szulc [1983] on the irrelevance of the functional form for the index number formula and on the unimportance of the choice of a fixed base or the use of chaining are confirmed for the Turvey data. However, we would still recommend the use of the chain principle whenever possible; for periods of time longer than the Turvey data time period, we would expect differences between fixed base and chained index numbers to be more dramatic.

It may be useful to stress once again the key to our suggested treatment of the seasonality problem in the context of measuring price level change: the seasonality problem vanishes if instead of making monthly or seasonal comparisons of prices, we compare prices for an entire year with the corresponding prices in a base year. Critics of our suggested method of addressing the seasonal adjustment problem might suggest that our proposal ignores one of the most important aspects of a seasonally adjusted series and that is to enable us to **forecast** a future annual inflation rate. However, if one refers back to Table 4, it

can be seen that it is quite easy to forecast say the annual 1971 rate (which was 1.11945) given the inflation rates for the first hald of 1971: the 1971 entries in the table exhibit a fairly steady growth rate with a slight acceleration in this rate. Thus the forecasting problem appears to be straightforward in the context of our suggested method of seasonal adjustment.

It may also be useful to compare our suggested method of seasonal adjustment with the methods that are used by most official statistical agencies. The starting point in "official" methods is usually the calculation of monthly price indexes using the Laspeyres formula for a fixed basket of goods that is somehow representative for the year as a whole. In our view, this first step is not very satisfactory. At the very least, monthly baskets should be introduced and then monthly indexes that compare prices of a month with the prices of the same month in a base year should be computed (e.g., see Table 6). Then "simple" annual indexes may be computed from these monthly indexes using a 12-month moving average with fixed weights equal to 1/12 (e.g., see Table 7). More "sophisticated" annual indexes may be computed from the monthly indexes by using more "sophisticated" annual indexes approximate our theoretically "most desirable" annual indexes (see Tables 4 or 5) very closely. We note that the prices in Table 6 differ from the corresponding prices in Tables 8 to about 1% accuracy. Thus the benefit from using our theoretically preferred weighting scheme (recall Table 8) over using the simple weighting scheme that generated the data in Table 7 is small but not negligible.

Our proposed method of seasonal adjustment should be useful in the producer context as well. The details are left to the reader.

To conclude: we have offered an economic justification for our suggested method for incorporating seasonal commodities into a cost-of-living index. In practice, our suggested method gives results which are reasonably close to the French method of taking 12-month moving averages of monthly indexes[12] (recall the construction of Table 7). In order to appreciate the enormous variation in the treatment of seasonal commodities in a CPI by official statistical agencies throughout the world, the paper by Turvey [1979] should be consulted.

Footnotes

* The author is greatly indebted to Bert Balk for helpful comments and to Denise Doiron for her excellent research assistance. This research was supported by Statistics Canada and the SSHRC of Canada. Neither institution is responsible for the views expressed here.

1 Turvey [1979; p.22] notes this ''solution'' along with several others. Turvey's discussion of the various possibilities is much more comprehensive than our brief discussion.

2 There are many additional time series methods papers in this volume that could be cited.

3 Diewert [1980; pp.501-503] outlines such an econometric procedure in the context of the new goods problem that will work equally well in the seasonal commodities context.

4 Allen [1975; pp.88-191] and Diewert [1980; p.507] advocate this principle.

5 Turvey [1979; p.22] notes this point. See Balk [1980a] [1980b] [1980c] [1981] for additional approaches.

6 Seasonal goods that are always consumed in zero amounts may be dropped from our list of goods, so for the Turvey data, we actually have only 60-16 = 44 goods.

7 $p \cdot x \equiv \Sigma_{n=1}^{44} p_n x_n$ denotes the inner product between the vectors p and x. Note that P_L does not actually depend on x^t while P_P does not depend on x^0.

8 Or to use Irving Fisher's [1922] terminology, deviations from circularity are small.

9 Our suggested seasonal adjustment procedure seems to be very close to the procedure used by France as reported by Turvey [1979].

10 This theory of bounds has been developed by Konüs [1924], Samuelson [1947], Malmquist [1953], Pollak [1971] and has been surveyed by Diewert [1983] for this conference.

11 We have factored out a common term δ_1.

12 For a description of the method, see Turvey [1979].

References

Allen, R.G.D. [1957], *Index Numbers in Theory and Practice*, London: Macmillan.

Balk, B.M. [1980a], *Seasonal Products in Agriculture and Horticulture and Methods for Computing Prices Indices,* Statistical Studies No. 24, The Hague: Netherlands Central Bureau of Statistics.

__________ [1980b], ''Seasonal Commodities and the Construction of Annual and Monthly Price Indexes'', *Statistiche Hefte* 21, pp.110-116.

__________ [1980c], ''A Method for Constructing Prices Indices for Seasonal Commodities'', *The Journal of the Royal Statistical Society* (Series A) 143, pp.68-75.

__________ [1981], ''A Simple Method for Constructing Price Indices for Seasonal Commodities'', *Statistische Hefte* 22, pp.72-78.

Diewert, W.E. [1978], ''Superlative Index Numbers and Consistency in Aggregation'', *Econometrica* 46, pp.883-900.

__________ [1980], ''Aggregation Problems in the Measurement of Capital'', pp.433-538 in *The Measurement of Capital*, Dan Usher (ed.), Chicago: The University of Chicago Press.

__________ [1983], ''The Theory of the Cost-of-Living Index and the Measurement of Welfare Change'', *Price Level Measurement: Proceeding From a Conference Sponsored by Statistics Canada.*

Fisher, Irving [1922], *The Making of Index Numbers*, Boston: Houghton-Mifflin.

Généreux, P.A. [1983], ''Impact of the Choice of Formulae on the Canadian Consumer Price Index'', *Price Level Measurement: Proceeding From a Conference Sponsored by Statistics Canada.*

Konüs, A.A. [1924], ''The Problem of the True Index of the Cost-of-Living'', translated in *Econometrica* 7, 1939, pp.10-29.

Kuiper, J. [1978], ''A Survey and Comparative Analysis of Various Methods of Seasonal Adjustment'', pp.59-76 in *Seasonal Analysis of Economic Time Series*, Proceedings of the Conference on the Seasonal Analysis of Economic Time Series, Arnold Zellner (ed.), Washington, D.C., September 9-10, 1976, Washington, D.C.: U.S. Government Printing Office.

Malmquist, S. [1953], ''Index Numbers and Indifference Surfaces, ''*Trabajos de Estadistica* 4, pp.209-242.

Pollak, R.A. [1971], ''The Theory of the Cost-of-Living Index'', Research Discussion Paper 11, Office of Prices and Living Conditions, U.S. Bureau of Labor Statistics, Washington, D.C.

Samuelson, P.A. [1947], ''The Theory of the Cost-of-Living Index'', Research Discussion Paper 11, Office of Prices and Living Conditions, U.S. Bureau of Labor Statistics, Washington, D.C.

__________ [1947], *Foundations of Economic Analysis*, Cambridge, Massachusetts: Harvard University Press.

Statistics Canada [1982], ''The Consumer Price Index Reference Paper: Concepts and Procedures'', Catalogue No. 62-553, Occasional, Ottawa.

Stone, R. [1956], *Quantity and Price Indexes in National Accounts*, Paris: The Organization for European Economic Co-operation.

Szulc, B.J. [1983], ''Linking Index Number Series'', *Price Level Measurement: Proceeding From a Conference Sponsored by Statistics Canada.*

Turvey, R. [1979], ''The Treatment of Seasonal Items in Consumer Price Indices'', *Bulletin of Labour Statistics* 4th quarter, International Labour Office, Geneva, pp.13-33.

SECTION IV

Theoretical Foundations of Output Indexes
Les bases théoriques des Indices de production

THE THEORY OF THE OUTPUT PRICE INDEX AND THE MEASUREMENT OF REAL OUTPUT CHANGE

*W.E. Diewert**
Department of Economics,
University of British Columbia

SUMMARY

The paper studies the problems involved in measuring price level change, real output change, and changes in efficiency in an open economy from the producer perspective as opposed to the consumer perspective. Section 1 of the paper discusses under what conditions the producer and consumer perspectives will yield the same numerical answers to the above measurement questions.

Section 2 develops the theory of the output price deflator. Various empirically implementable approximations to the theoretical index are suggested.

Section 3 develops various theoretical concepts of real output indexes from the perspective of producer theory. Again, various approximations to the theoretical indexes are suggested.

Section 4 develops the theory of productivity indexes. These indexes indicate how much an economy has grown due to increases in efficiency (or changes in the terms of trade facing the economy over two periods), adjusting for changes in inputs utilized by the economy during the two periods. In Section 5, the paper suggests various techniques which may be used in order to determine what proportion of an economy's productivity growth is due to changes in the external economic environment facing the economy during the two periods (changes in the terms of trade; i.e., changes in the prices of exports relative to imports).

Section 6 discusses some practical problems with the theoretical indexes suggested in the first five sections of the paper.

RÉSUMÉ

La communication aborde les problèmes que comporte la mesure des variations de prix, des changements dans la production et d'efficacité dans une économie ouverte dans la perspective des producteurs plutôt que celle des consommateurs.

La première section met l'accent sur les facteurs qui peuvent conduire à des réponses différentes face aux problèmes de mesure qui précèdent, lorsqu'on se place dans la perspective de la théorie du producteur plutôt que dans celle de la théorie du consommateur. Ainsi, cette section passe en revue et met à jour les vieilles controverses, qui datent de Hicks et de Samuelson, sur l'approche à adopter à l'égard de la comptabilité du revenu national.

La section 2 passe en revue et étend la théorie de l'indice des prix de la production (ou du déflateur du PNB), qu'on attribue à Fisher et Shell, à Samuelson et Swamy et à Archibald. Nous proposons également des approximations empiriquement applicables des divers indices de prix théoriques.

La section 3 passe en revue et étend les divers concepts d'indices de la production réelle qui ont été proposés. Il y a trois grandes approches qui ont été proposées dans les textes économiques: (i) l'approche de Fisher et Shell (qui divise le ratio nominal du PNB pour deux périodes par l'indice correspondant des prix de la production de Fisher-Shell, (ii) l'approche de Samuelson-Swamy et Sato, et (iii) l'approche de Malmquist, Bergson et Moorsteen. Nous développons, à la section 3, des conditions nécessaires et suffisantes d'équivalence entre les trois approches, ainsi que diverses approximations empiriquement applicables des indices théoriques.

Les indices de la production étudiés à la section 3 indiquent la croissance de production dans la période en cours par rapport à la période de base. Cette croissance de production pourrait être attribuable à: (i) une croissance des intrants, (ii) des accroissements d'efficacité technique (progrès technique) et (iii) des changements des termes d'échange (c.-à-d. des

changements des prix des exportations par rapport aux importations). À la section 4, nous isolons la contribution de la croissance de la production ou de la croissance des intrants et étudions des indices d'efficacité. Ces indices s'appellent parfois des indices de la productivité totale des facteurs.

À la section 5, nous isolons l'apport de la croissance de la production des changements des termes d'échange. Nous proposons divers facteurs théoriques de correction des termes d'échange et tirons des approximations empiriques de ces indices théoriques.

La section 6 traite de certains problèmes pratiques que posent les résultats théoriques obtenus plus tôt dans la communication. Ces difficultés sont: (i) l'existence de trop nombreux biens, (ii) l'existence d'intrants durables (composantes du stock de capital), (iii) l'existence de taxes et d'impôts et (iv) l'existence d'un comportement non concurrentiel de la part des producteurs.

1. Introduction

This paper is concerned with the measurement of (i) welfare change and (ii) inflation in an open economy.

The above measurement problems are trivial in a one person, one variable good economy: if we denote the (variable) output of the economy in period t by $y^t > 0$ and the corresponding price of output by $p^t > 0$, then we may define the rates of welfare change and of inflation going from period 0 to period 1 as y^1/y^0 and p^1/p^0 respectively. Note that we may also interpret y^1/y^0 as the rate of growth in real output or consumption.

As soon as we leave the above hypothetical world, we have two different methods for measuring (i) and (ii): a consumer theory oriented method which is based on the assumption that consumers maximize their own welfare or utility functions subject to budget constraints,[1] and a producer theory oriented method which is based on the assumption that producers maximize profits subject to their technological constraints.[2]

Hicks [1975; p.317] (see also Hicks [1939, 1940]) notes that the two methods of measurement can lead to the same answers in the many goods case provided that the following assumptions are satisfied:

(i) the economy is closed, so that there is no international trade,

(ii) there are no goods that last longer than the period under consideration, so that there are no consumer durables or investment goods,

(iii) all consumers in the economy have identical preferences,

(iv) all consumers have the same income,

(v) the technology is well-behaved with no increasing returns to scale and no externalities,

(vi) there is price taking behaviour on the part of all consumers and producers,

(vii) there are no taxes (other than lump sum transfers) and

(viii) there is no government sector, no public goods, no nonprofit organizations so that gross national product coincides with private (profit maximizing) production.[3]

Under the above restrictive hypotheses, the consumer theory based true cost-of-living index will closely approximate the producer theory based output price index that we will define below in Section 2; in fact, under these conditions, the consumer based Paasche and Laspeyres price indexes will equal the corresponding producer based Paasche and Laspeyres price indexes. Moreover, the growth in consumer welfare will closely approximate the growth of real output under the above restrictive assumptions.

Since the hypotheses (i)-(viii) above are not even approximately satisfied, it is clear that the consumer and producer theory based measurement methods will generally lead to different answers. Hence we must study each approach separately.

We looked at the consumer theory methods of measurement in Diewert [1983] in some detail. Our main conclusions there were that the consumer theory oriented method works reasonably well from the viewpoint of measuring inflation, but the measurement of welfare change from the consumer point of view works well only if we measure welfare change for reasonably homogeneous groups of consumers or if we are able to make rather specific assumptions about the relative welfare of different consumers.

In this paper, we shall focus on the measurement of output growth and inflation from the producer theory perspective. Thus in Section 2, we present the Fisher-Shell [1972], Samuelson and Swamy [1974] and Archibald [1977] theory of the output price index (or the nominal GNP deflator as it is sometimes called). In Section 3, we present the corresponding theory of the measurement of real output change while Section 4 presents the closely related theory of (total factor) productivity indexes. In Section 5, we indicate a method for measuring the contribution of changes in a country's terms of trade to output price inflation. Section 6 lists the main difficulties with our methodology. Section 7 concludes.

However, before we jump into the technical aspects of measurement from the producer theory perspective, it may be useful to provide some motivation for measuring growth and inflation from the producer rather than the consumer perspective.

The first point to note is that the concept of real output growth seems to be much more acceptable to the public at large than the corresponding consumer oriented concept of welfare change. This is perhaps not surprising, given the difficulty in deciding how different consumer groups should be weighted when constructing the welfare index. Thus the popular literature is filled with references to a country's growth in real output (and to cross country comparisons of per capita real output). Hence it appears to be useful to provide a theoretical justification for a real output index and some practical guidance on how to best approximate the theoretical construction.

The second point we wish to note is due to Hicks [1939, 1940, 1975] and Samuelson [1950; pp.12-13]: if more of any good can be produced by a country during period 1 compared to period 0 without reducing the production of any other good (or increasing the utilization of variable inputs such as labour), then the country's production possibilities

set has shifted outward (or to use alternative terminology, there has been technical progress going from period 0 to 1), and there is an increase in society's potential welfare; i.e., with the appropriate tax and transfer policy, the government **could** ensure that the increase in productive efficiency could be translated into a welfare increase for each homogeneous consumer group in the economy. Thus if inputs into the economy remain constant, growth in an economy's real output index could indicate a potential increase in consumer welfare.

The argument in point two above hypothesized that the utilization of variable inputs remain constant during the two periods for which comparisons are being made. Of course, usually variable inputs into the private production sector of an economy will not remain constant. However, let us follow the suggestion of Samuelson [1950; p.23] and treat labour and other variable inputs as negative outputs. The counterpart to the real output index now becomes a (total factor) **productivity** index. Outward shifts in a country's production possibilities set will lead to a greater than unity productivity index, as we shall see in Section 4 below.

In the analysis which follows, prices are to be regarded as prices from the viewpoint of producer theory; i.e., the price of an output produced by a producer is the price excluding commodity and excise taxes but including subsidies, whereas the price of an input used by a producer is the price including any relevant commodity, excise, sales or income taxes on the input. The relevant prices are the revenues actually received by the producer for one unit of an output and costs paid out for one unit of an input.

Finally, we note that the analysis which follows is also applicable to a single firm, or to a collection of price taking competitive firms comprising a sector of the economy, as well as to the entire private sector of the economy.[4]

2. The Theory of the Output Price Deflator

Suppose that the set of feasible outputs and variable inputs of the private production section of an economy during period t is denoted by S^t, a subset of M + N dimensional space. Thus if $(y,v) \epsilon S^t$, then the nonnegative M dimensional vector of outputs

$y \equiv (y_1,\ldots,y_M) \geq 0_M$ is producible by the economy if producers can utilize the nonnegative N dimensional vector of inputs $v \equiv (v_1,\ldots,v_N) \geq 0_N$.

Let $p \equiv (p_1,\ldots,p_M) >> 0_M$, denote a positive vector of output prices that producers in the economy face. Define the economy's (nominal) **national product function**[5] for $p >> 0_M$, $v \geq 0_N$ by

$$\pi(p,v) \equiv \max_y \left\{p \cdot y: \quad (y,v)\epsilon S^t\right\} \tag{1}$$

where $p \cdot y \equiv \Sigma_{m=1}^{M} p_m y_m$ denotes the inner product between the vectors p and y. Thus $\pi(p,v)$ is the maximum value of outputs that the economy can produce in period t given that the vector of variable inputs v is available and given that producers face the same output price vector p.

We follow the example of Fisher and Shell [1972] (see also Samuelson and Swamy [1974; pp.588-592], Archibald [1977; pp.60-61] and Diewert [1980; p.461] and define the economy's **output price index** between periods 0 and 1 using the technology of period t as

$$P^t(p^0,p^1,v) \equiv \pi^t(p^1,v)/\pi^t(p^0,v); \qquad t = 0 \text{ or } 1 \tag{2}$$

where $p^t >> 0_M$ is the vector of output prices the economy faces in period t for t = 0, 1 and $v > 0_N$ is a nonnegative nonzero reference vector of inputs. Note that in definition (2), the technology and the level of input utilization is held constant; the only variable that varies is p, the output price vector that the economy faces. We further note that if $M = 1$ so that there is only one output, then (2) collapses down to p_1^1/p_1^0, the output price ratio for the single output.

The reader who is familiar with the theory of the cost-of-living index (e.g., see Konüs [1924], Pollak [1971] or Diewert [1983]) should note the analogy of the output price index to the cost-of-living index. In the theory of the consumer's cost-of-living index, the producer's national product function or value added function $\pi(p,v)$ is replaced by the consumer's minimum cost or expenditure function $C(p,u)$, the t and v which occur in (2) are replaced with a single **scalar** u which indexes the reference standard of living (or indifference surface or utility level), and the producer price vectors which occur in (2) are replaced

by consumer price vectors. As we noted earlier, the producer and consumer price vectors will not be identical due to different domains of definition (i.e., the existence of internationally traded intermediate inputs, capital goods, consumer durables, etc.) and the existence of taxes. The reader should also note that there are a wide variety of Fisher-Shell output price indexes of the form (2), depending, on which (t,v) reference technology and input utilization vector we choose.

Usually, we are interested in two special cases of the general definition (2), namely:

$$P^0(p^0,p^1,v^0) \qquad \text{(period 0 technology and inputs) and} \tag{3}$$

$$P^1(p^0,p^1,v^1) \qquad \text{(period 1 technology and inputs)} \tag{4}$$

where $v^t > 0_N$ is the actual vector of inputs utilized by the economy during period t for t = 0, 1. For a geometric interpretation of the indexes (3) and (4), see Fisher and Shell [1972; p.52].

Let us call the index defined by (3) the **Laspeyres Fisher-Shell** output price index and the index defined by (4) the **Paasche Fisher-Shell** output price index.

We now prove two theorems that provide some (nonparametric) bounds for price indexes of the form (2), (3) or (4).

Theorem 1: Let the positive output price vectors $p^0 >> 0_N$ and $p^1 >> 0_M$ be given along with the reference input vector $v > 0_N$. Suppose the technology set S^t has the following property: $\left\{y\colon (y,v)\epsilon S^t\right\}$ is a nonempty, closed and bounded subset of the nonnegative orthant in R^M (Euclidean M dimensional space) that does not coincide with 0_M. Then the output price index defined by (2) satisfies the following equalities:

$$\min_m \left\{p_m^1/p_m^0\colon \ m=1,\ldots,M\right\} \le \ P^t(p^0,p^1,v) \le \max_m \left\{p_m^1/p_m^0\colon \ m=1,\ldots,M\right\} \tag{5}$$

Proof: Our assumptions on S^t imply the existence of y^0,y^1 such that

$$\pi^t(p^1,v) \equiv \max_y \left\{p^1 \cdot y: \ (y,v)\epsilon S^t\right\} = p^1 \cdot y^1 > 0 \text{ and} \qquad (6)$$

$$\pi^t(p^0,v) \equiv \max_y \left\{p^0 \cdot y: \ (y,v)\epsilon S^t\right\} = p^0 \cdot y^0 > 0. \qquad (7)$$

It can be shown that

$$\left\{y: \ (y,v)\epsilon S^t\right\} \text{ is a subset of } \left\{y: \ p^0 \cdot y \leq p^0 \cdot y^0, \ y \geq 0_M\right\} \qquad (8)$$

since if (8) were not true, then the equality in (7) would be false. Similarly, we have (using (6))

$$\left\{y: \ (y,v)\epsilon S^t\right\} \text{ is a subset of } \left\{y: p^1 \cdot y \leq p^1 \cdot y^1, \ y \geq 0_M\right\} . \qquad (9)$$

Jsing definitions (1) and (2), we have

$${}^t(p^0,p^1,v) \equiv \max_y \left\{p^1 \cdot y: \ (y,v)\epsilon S^t\right\} / \pi^t(p^0,v)$$

$$= \max_y \left\{p^1 \cdot y: \ (y,v)\epsilon S^t\right\} / p^0 \cdot y^0 \qquad \text{using (7)}$$

$$\leq \max_y \left\{p^1 \cdot y: \ p^0 \cdot y \leq p^0 \cdot y^0, \ y \geq 0_M\right\} / p^0 \cdot y^0 \qquad \text{using (8)}$$

i.e., the maximum of a function over a bigger set cannot decrease

$$= \max_i \left\{p_i^1 \ p^0 \cdot y^0 / p_i^0 : i = 1,\ldots,M\right\} / p^0 \cdot y^0$$

upon solving the linear programming problem

$$= \max_i \left\{p_i^1 / p_i^0 : i = 1,\ldots,M\right\}$$

hich establishes the right-hand inequality in (5). The left-hand inequality follows in a milar manner, starting with the equation $P^t(p^0,p^1,v) = p^1 \cdot y^1 / \max_y \left\{p^0 \cdot y: (y,v)\epsilon S^t\right\}$ nd using (9), etc. Q.E.D.

The method of proof used above is essentially due to Samuelson [1947; p.159].

heorem 2: (Fisher and Shell [1972; pp.57-58], Archibald [1977; p.66]) Assume p^0,p^1,v nd S^t satisfy the same regularity conditions as those listed in Theorem 1 for $t = 0, 1$.

The special output price indexes defined by (3) and (4) satisfy the following inequalities:

$$P^0(p^0,p^1,v^0) \geq p^1 \cdot y^0 / p^0 \cdot y^0 \equiv \text{the Laspeyres price index, } P_L, \qquad (10)$$

$$P^1(p^0,p^1,v^1) \leq p^1 \cdot y^1 / p^0 \cdot y^1 \equiv \text{the Paasche price index, } P_P. \qquad (11)$$

Proof: $P^0(p^0,p^1,v^0) \equiv \pi^0(p^1,v^0)/\pi^0(p^0,v^0)$

$$\equiv \max_y \left\{ p^1 \cdot y: (y,v^0) \epsilon S^0 \right\} / p^0 \cdot y^0$$

$$\geq p^1 \cdot y^0 / p^0 \cdot y^0$$

since y^0 is feasible for the maximization problem. The proof of (11) is similar. Q.E.D.

We note that the method of proof used in the above theorem (which might be called "revealed production theory") dates back to Hicks [1940] and Samuelson [1950].

Corollary 2.1: P^0 and P^1 satisfy the following observable bounds:

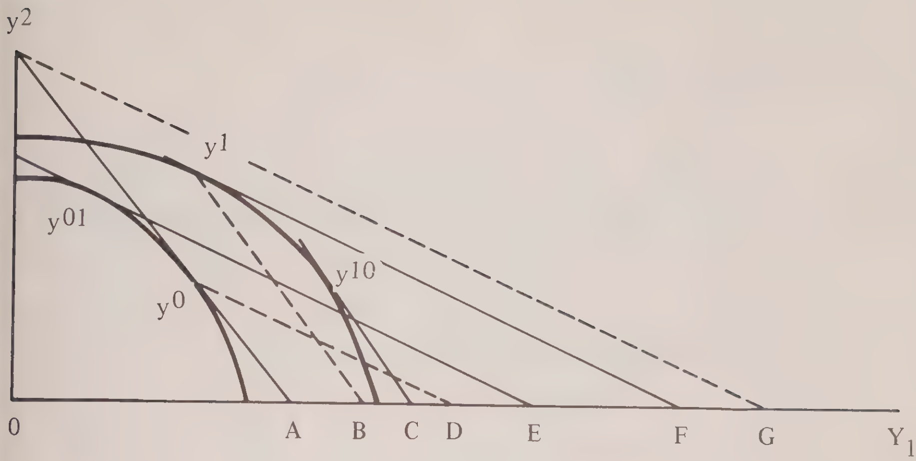
$$P_L \equiv p^1 \cdot y^0 / p^0 \cdot y^0 \leq P^0(p^0,p^1,v^0) \leq \max_i \left\{ p_i^1 / p_i^0: \ i=1,\dots,M \right\}; \qquad (12)$$

$$\min_i \left\{ p_i^1 / p_i^0: \ i=1,\dots,M \right\} \leq P^1(p^0,p^1,v^1) \leq p^1 \cdot y^1 / p^0 \cdot y^1 \equiv P_P. \qquad (13)$$

The corollary follows by combining Theorems 1 and 2. This corollary is similar to a corollary noted by Pollak [1971] in the context of consumer theory.

We illustrate the above bounds for the case of two outputs ($M=2$) in Figure 1. The inner curved line encloses the set of period 0 feasible outputs, while the more northeast curved line encloses the set of period 1 feasible outputs. The vector $y^0 \equiv (y_1^0,y_2^0)$ is the observed period 0 output vector while $y^1 \equiv (y_1^1,y_2^1)$ is the observed period 1 output vector. If producers were faced with period 1 (0) prices and had period 0 (1) inputs and technology at their disposal, then they would collectively choose the profit maximizing output vector

Figure 1



y^{01} (y^{10}) which is unobservable. By the definition of the (unobservable) Laspeyres Fisher-Shell output price index (3), $P^0(p^0,p^1,v^0) \equiv \pi^0(p^1,v^0)/\ \pi^0(p^0,v^0) = p^1 \cdot y^{01}/p^0 \cdot y^0$. The (observable) Laspeyres price index is $P_L \equiv p^1 \cdot y^0/p^0 \cdot y^0$. Hence $P^0(p^0,p^1,v^0)/P_L = p^1 \cdot y^{01}/p^1 \cdot y^0 = OE/OD > 1$ using Figure 1. Also $P^0(p^0,p^1,v^0)/\max_i \left\{p_i^1/p_i^0 : i=1,2\right\} = OE/OG < 1$ using Figure 1. These inequalities correspond to the bounds in (12). Similarly, $P^1(p^0,p^1,v^1) \equiv \pi^1(p^1,v^1)/\pi^1(p^0,v^1) = p^1 \cdot y^1/p^0 \cdot y^{10}$, $P_P \equiv p^1 \cdot y^1/p^0 \cdot y^1$, $P^1(p^0,p^1,v^1)/P_P = p^0 \cdot y^1/p^0 \cdot y^{10} = OB/OC < 1$ and $P^1(p^0,p^1,v^1)/\min_i \left\{p_i^1/p_i^0 : i=1,2\right\} = OF/OC > 1$. The last two inequalities correspond to the bounds in (13).

Unfortunately, the bounds in (12) and (13) above may often be rather wide in practical situations. We conclude this section by outlining two alternative strategies for obtaining tighter bounds. In the first strategy, we attempt to choose a convenient reference vector v and in the second strategy, we assume a convenient functional form for π^t.

The dependence of $\pi^t(p,v)$ on t causes us some problems initially: we must first define a national product function for the economy which is based on an **average** of the technology sets S^0 and S^1 for the period 0 and 1. Thus for $0 \le t \le 1$, $p >> 0_M$ and $v \ge 0_N$, define the following **average national product function**:

$$\pi(p,v,t) \equiv \max_y \left\{p \cdot y: \ (y,v,)\epsilon \ [(1-t)S^0 + tS^1]\right\}. \tag{14}$$

Note that

$$\pi(p,v,0) = \pi^0(p,v) \text{ and } \pi(p,v,1) = \pi^1(p,v). \tag{15}$$

Thus $\pi(p,v,t)$ is the national product function that is generated by a weighted average of the two technology sets where S^0 gets the weights $(1\text{-}t) \geq 0$ and S^1 gets the weight $t \geq 0$. The following theorem shows that there is an average reference input vector of the form $(1\text{-}t)v^0 + tv^1$ and an average technology such that the Paasche and Laspeyres output price indexes bound a theoretical output price index of the form

$$P(p^0,p^1,v,t) \equiv \pi(p^1,v,t)/\pi(p^0,v,t). \tag{16}$$

Theorem 3: Suppose the economy faces the positive output price vectors $p^0 >> 0_M$ and $p^1 >> 0_M$ and utilizes the positive input vectors $v^0 >> 0_N$ and $v^1 >> 0_N$ during periods 0 and 1. Suppose in addition that the period 0 technology set S^0 and the period 1 technology set S^1 are sufficiently well-behaved[6] so that $\pi(p^1,(1\text{-}t)v^0 + tv^1,t)$ and $\pi(p^0,(1\text{-}t)v^0 + tv^1,t)$ are well-defined by (14) and continuous positive functions of t for $0 \leq t \leq 1$. Then there exists t^* such that $0 \leq t^* \leq 1$ and the output price index $P(p^0,p^1,(1\text{-}t^*)v^0 + t^*v^1,t^*)$ lies between the Paasche and Laspeyres output price indexes, i.e., if $P_L \leq P_P$, we have (17), while if $P_L \geq P_P$, we have (18):

$$P_L \equiv p^1 \cdot y^0/p^0 \cdot y^0 \leq P(p^0,p^1,(1\text{-}t^*)v^0 + t^*v^1,t^*) \leq p^1 \cdot y^1/p^0 \cdot y^1 \equiv P_P \text{ or} \tag{17}$$

$$P_P \leq P(p^0,p^1,(1\text{-}t^*)v^0 + t^*v^1,t^*) \leq P_L. \tag{18}$$

Proof: Define $h(t) \equiv P(p^0,p^1,(1\text{-}t)v^0 + tv^1,t) = \pi(p^1,(1\text{-}t)v^0 + tv^1,t)/\pi(p^0,(1\text{-}t)v^0 + tv^1,t)$, a continuous function for $0 \leq t \leq 1$. Note that $h(0) \equiv \pi(p^1,v^0,0)/\pi(p^0,v^0,0) = \pi^0(p^1,v^0)/\pi^0(p^0,v^0)$ using (15), and $h(1) \equiv \pi(p^1,v^1,1)/\pi(p^0,v^1,1) = \pi^1(p^1,v^1)/\pi^1(p^0,v^1)$ using (15) again. There are 24 possible *a priori* inequality relations that are possible between the four numbers h(0), h(1), P_L and P_P. However, Theorem 2 implies $h(0) = P^0(p^0,p^1,v^0) \geq P_L$ and $h(1) = P^1(p^0,p^1,v^1) \leq P_P$. This means that there are only six possible inequalities between the four numbers: (1) $h(0) \geq P_L \geq P_P \geq h(1)$, (2) $h(0) \geq$

$P_P \geq P_L \geq h(1)$, (3) $h(0) \geq P_P \geq h(1) \geq P_L$, (4) $P_P \geq h(0) \geq P_L \geq h(1)$, (5) $P_P \geq h(1) \geq h(0) \geq P_L$ and (6) $P_P \geq h(0) \geq h(1) \geq P_L$. Since h(t) is continuous over $0 \leq t \leq 1$, it assumes all intermediate values and hence there exists $0 \leq t^* \leq 1$ such that (17) is true if case (1) to (5) holds or (18) is true if case (6) holds. Q.E.D.

The method of proof used above is the same as that used by Konüs [1924] in the consumer context.

Theorem 3 is a very satisfactory result from an empirical point of view since P_L and P_P will generally be quite close to each other for data that comes from adjacent time periods.[7] It is also a very satisfactory result from a theoretical point of view, since we have made no restrictive assumptions about the functional form of π^0 or π^1; we do not even require the differentiability of the national product functions π^t.

The following theorem requires that the functional forms for the national product functions π^0 and π^1 be of the following **translog** functional form: for t = 0, 1

$$\ln\pi^t(p,v) \equiv \alpha_0^t + \sum_{n=1}^{N} \beta_n^t \ln v_n + \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} \beta_{ij}^t \ln v_i \ln v_j \tag{19}$$

$$+ \sum_{m=1}^{M} \alpha_m^t \ln p_m + \frac{1}{2}\sum_{i=1}^{M}\sum_{j=1}^{M} \alpha_{ij}^t \ln p_i \ln p_j + \sum_{m=1}^{M}\sum_{n=1}^{N} \gamma_{mn}^t \ln p_m \ln v_n$$

where $\beta_{ij}^t = \beta_{ji}^t$ and $\alpha_{ij}^t = \alpha_{ji}^t$ for all i and j. The functional form on the right-hand side of (19) can approximate an arbitrary twice continuously differentiable function, f(p,v) say, to the second order; i.e., the translog functional form is flexible to use Diewert's [1974; p.113] terminology.

Theorem 4 below also requires the definition of the Törnqvist or **translog output price index P_T:**

$$P_T(p^0,p^1,y^0,y^1) \equiv \prod_{i=1}^{M} (p_i^1/p_i^0)^{[\frac{1}{2}s_i^0 + \frac{1}{2}s_i^1]} \tag{20}$$

where $s_i^t \equiv p_i^t y_i^t / p^t \cdot y^t$ for $t = 0,1$ and $i = 1,\ldots,M$. Note that P_T is like P_L or P_P: given the observable price and output vectors p^t and y^t for periods $t = 0$ and 1, we can numerically compute $P_T(p^0,p^1,y^0,y^1)$.

Theorem 4: Make the following assumptions: (i) the national product function for the economy in period t, π^t, has the translog functional form defined by (19) for $t = 0$ and 1, (ii) the economy faces the positive output price vector $p^t >> 0_M$ and has available the positive input vector $v^t >> O_N$ during periods $t = 0,1$, (iii) $y^t > 0_M$ is the revenue maximizing output vector for the economy during period t; i.e., $p^t \cdot y^t = \pi^t(p^t,v^t)$ for $t = 0,1$, and (iv) $\alpha_{ij}^0 = \alpha_{ij}^1$ for $1 \leq i, j \leq M$; i.e., the coefficients for the quadratic terms in the logarithms of output prices are identical for the two national product functions π^0 and π^1 (but the other coefficients may be completely different between periods 0 and 1). Then a geometric average of the theoretical output price indexes defined by (3) and (4) above is exactly equal to the translog price index P_T defined by (20); i.e.,

$$[P^0(p^0,p^1,v^0)P^1(p^0,p^1,v^1)]^{1/2} = P_T(p^0,p^1,y^0,y^1). \qquad (21)$$

The proof of the above theorem is omitted since it is rather technical and it is virtually identical to the proof of Theorem 1 in Caves, Christensen and Diewert [1982; p.1398] with some minor changes.

Are the results of Theorems 3 and 4 inconsistent? Of course, the answer is no since each theorem makes different assumptions. However, from an empirical point of view, we will generally find that P_L, P_P and P_T are all rather close numerically; in fact Diewert [1978] showed that P_L, P_P and P_T will numerically approximate each other to the **first** order when each function is evaluated to a point where $p^0 = p^1$ and $y^0 = y^1$. If we take a symmetric average of P_L and P_P such as Fisher's [1922] ideal price index, $P_F(p^0,p^1,y^0,y^1) \equiv [P_L P_P]^{1/2}$, then Diewert [1978] also showed that P_F and P_T would numerically approximate each other to the **second** order. Numerical experiments with real economic data typically show the P_L, P_P, P_F and P_T coincide to three significant figures, while P_F and P_T coincide to four significant figures. Thus Theorems 3 and 4 lead to virtually identical numerical results.

Our conclusion at this stage is that it is possible to give a reasonably firm theoretical foundation (that leads to reasonable empirical results) for the concept of an output price index. Of course, the two theoretical output price indexes that we are able to calculate or approximately calculate, $[P^0(p^0,p^1,v^0)P^1(p^0,p^1,v^1)]^{1/2}$ (Theorem 4) and $P(p^0,p^1,(1-t^*)v^0 + t^*v^1,t^*)$ (Theorem 3), are only two output price indexes out of an infinite number of possible theoretical indexes. However, perhaps these two indexes will suffice for most practical purposes.

3. The Theory of Output Indexes

One can distinguish at least three different approaches to the construction of a real output index based on production theory:

(i) The **Fisher-Shell** [1972; p.53] **Approach**. In this approach, once the "correct" output price index has been found, then the corresponding "correct" index of real output is obtained residually, by dividing the nominal value of output ratio for the two periods, $p^1 \cdot y^1/p^0 \cdot y^0$, by the output price index. Thus if we take the correct price index to be $P^t(p^0,p^1,v)$ defined by (2) above, then the **Fisher-Shell output index** is defined as[8]

$$Q^t_{FS}(p^0,p^1,v,y^0,y^1) \equiv p^1 \cdot y^1/p \cdot y^0 P^t(p^0,p^1,v). \qquad (22)$$

(ii) the **Samuelson-Swamy** [1974; p.588]- **Sato** [1976; p.438] **Approach**. In this approach, the nominal value added function $\pi^t(p,v)$ defined above by (1), is utilized: for any reference output price vector $p >> 0_M$, defined the **Samuelson-Swamy-Sato output index** by[9]

$$Q_{SSS}(v^0,v^1,p) \equiv \pi^1(p,v^1)/\pi^0(p,v^0) \qquad (23)$$

$$= [\pi^1(p,v^1)/\pi^0(p,v^1)][\pi^0(p,v^1)/\pi^0(p,v^0)].$$

Note that the index (23) combines the effects of technical progress and input growth; i.e., the term $\pi^1(p,v^1)/\pi^0(p,v^1)$ represents the percentage increase in output (valued at the

reference output price vector p) that can be produced by the period 1 technology compared to the period 0 technology using the period 1 vector of inputs v^1, while the term $\pi^0(p,v^1)/\pi^0(p,v^0)$ represents the percentage increase in output that could be produced using the period 1 input vector v^1 compared to the period 0 input vector v^0 (using the base period technology and the reference output prices p). If there is no technological change so that $S^0 = S^1 \equiv S$, then $\pi^t(p,v) \equiv \max_y \left\{p\cdot y: (y,v)\epsilon S^t\right\} = \pi(p,v)$, and the Fisher-Shell output index may be written as a product of Samuelson-Swamy-Sato output indexes; i.e., we have for $v >> 0_N$,

$$Q_{FS}(p^0,p^1,v,y^0,y^1) = [\pi(p^1,v^1)/\pi(p^0,v^0)]/P^t(p^0,p^1,v) \tag{24}$$

$$= [\pi(p^1,v^1)/\pi(p^0,v^0)]/[\pi(p^1,v)/\pi(p^0,v)] \quad \text{(using(2)}$$

$$= [\pi(p^1,v^1)/\pi(p^1,v)][\pi(p^0,v)/\pi(p^0,v^0)]$$

$$= Q_{SSS}(v^0,v,p^0)Q_{SSS}(v,v,^1,p^1).$$

(iii) The **Malmquist** [1953]-**Bergson** [1961]- **Moorsteen** [1961] **Approach**. In this approach, it is first necessary to define the **output deflation function** d^t for the period t technology: for a positive output vector $y >> 0_M$ and a positive reference input vector $v >> 0_N$, define[10]

$$d^t(y,v) \equiv \min_{\delta>0} \left\{\delta: \quad (y/\delta,v)\epsilon S^t\right\}. \tag{25}$$

Thus $d^t(y,v)$ tells us by what proportion we have to deflate the output vector y so that the deflated output vector and the reference input vector v are just on the frontier of the period t production possibilities set S^t. The **Malmquist-Bergson-Moorsteen**[11] **output index** using period t technology and the reference input vector $v >> 0_N$ for the positive output vectors for periods 0 and 1, $y^0 >> 0_M$ and $y^1 >> 0_M$, may now be defined as

$$Q^t(y^0,y^1,v) \equiv d^t(y^1,v)/d^t(y^0,v). \tag{26}$$

The main property which (26) has (and (22) and (23) do not have in general) is the following homogeneity property: for every $v >> 0_N$, $y^0 >> 0_M$ and $\lambda > 0$, if $y^1 = \lambda y^0$ so that the period 1 output vector is proportional to the period 0 output vector, then $Q^t(y^0,y^1,v)$ equals λ, the common proportionality factor; i.e., for $y^0 >> 0_M$, $v >> 0_N$ and $\lambda > 0$,

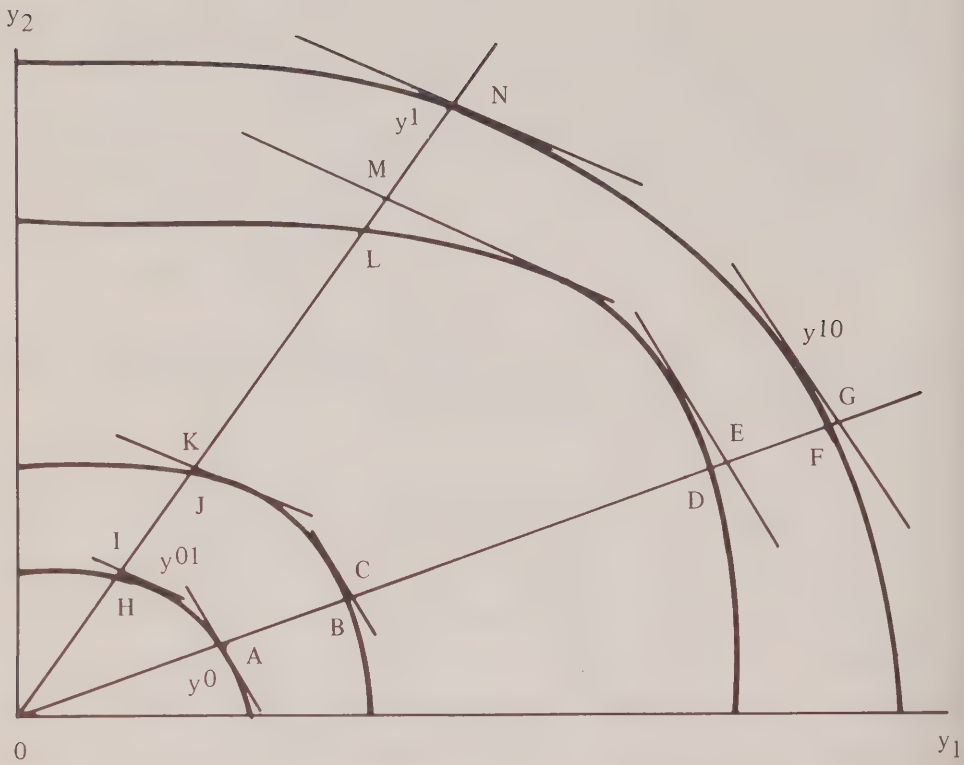
$$Q^t(y^0,\lambda y^0,v) = \lambda. \tag{27}$$

In the two output case, various special cases of the three classes of quantity indexes defined by (22), (23) and (26) may be illustrated using Figure 2.

Note that A in Figure 2 is counterpart to y^0 in Figure 1 and N corresponds to y^1. The point A is on the output production possibilities set $\{y: (y,v^0)\epsilon S^0\}$ that corresponds to the period 0 technology set S^0 and input vector v^0. The point N is on the output production possibilities set $\{y: (y,v^1)\epsilon S^1\}$ that corresponds to the period 1 technology set S^1 and input vector v^1. The point L is on the output production possibilities set $\{y: (y,v^1)\epsilon S^0\}$ that corresponds to the period 0 technology set and the period 1 input vector while the point J is on the set $\{y: (y,v^0)\epsilon S^1\}$ that corresponds to the period 1 technology set and the period 0 input vector. The diagram shows that there was both technological progress and input growth going from period 0 to period 1. There are two ways of going from the period 0 output production possibilities set, the smallest set, to the period 1 set, the largest set: (i) via the set containing J (technical progress first, then input growth) or (ii) via the set containing L (input growth first, then technical change).

Recall that $P^0(p^0,p^1,v^0) \equiv \pi^0(p^1,v^0)/\ \pi^0(p^0,v^0) = p^1\cdot y^{01}/p^0\cdot y^0$ and $Q^0_{FS}(p^0,p^1,v^0,y^0,y^1) \equiv [p^1\cdot y^1/p^0\cdot y^0]/[\pi^0(p^1,v^0)/\pi^0(p^0,v^0)] = p^1\cdot y^1/p^1\cdot y^{01} =$ ON/OI using Figure 2. Similarly, $Q^1_{FS}(p^0,p^1,v^1,y^0,y^1) \equiv [p^1\cdot y^1/p^0\cdot y^0]/[\ \pi^1(p^1,v^1)/\pi^1(p^0,v^1)] = [p^1\cdot y^1/p^0\cdot y^0]/[p^1\cdot y^1/p^0\cdot y^{10}] = p^0\cdot y^{10}/p^0\cdot y^0 =$ OG/OA using Figure 2. Also $Q^0_{FS}(p^0,p^1,v^1,y^0,y^1) \equiv [p^1\cdot y^1/p^0\cdot y^0]/[\ \pi^0(p^1,v^1)/\pi^0(p^0,v^1)] = [p^1\cdot y^1/\pi^0(p^1,v^1)]. [\pi^0(p^0,v^1)/p^0\cdot y^0] =$ [0N/0M][0E/0A]. Similarly, $Q^1_{FS}(p^0,p^1,v^0,y^0,y^1) \equiv [p^1\cdot y^1/p^0\cdot y^0]/ [\pi^1(p^1,v^0)/\pi^1(p^0,v^0)] = [\pi^1(p^0,v^0)/\pi^0(p^0,v^0)][\pi^1(p^1,v^1)/\pi^1(p^1,v^0)] =$ [0C/0A][0N/0K].

Figure 2



Turning now to the output indexes defined by (23), we have $Q_{SSS}(v^0,v^1,p^0) \equiv \pi^1(p^0,v^1)/\pi^0(p^0,v^0) = [p^0 \cdot y^{10}/p^0 \cdot y^0] = 0G/0A = Q^1_{FS}(p^0,p^1,v^1,y^0,y^1)$ in general and $Q_{SSS}(v^0,v^1,p^1) \equiv \pi^1(p^1,v^1)/\pi^0(p^1,v^0) = Q^0_{FS}(p^0,p^1,v^0,y^0,y^1) = 0N/0I$. Thus if we choose the reference prices in the Samuelson-Swamy-Sato output index to be p^0 or p^1, we obtain special cases of the Fisher-Shell output indexes, (an equivalence which does not hold for arbitrary reference price vectors except when the technology satisfies certain restrictive separability assumptions; see Theorem 5 below).

Finally, we turn to the Malmquist type output indexes defined by (26). The two simplest cases are $Q^0(y^0,y^1,v^0) \equiv d^0(y^1,v^0)/d^0(y^0,v^0) = d^0(y^1,v^0) = 0N/0H$ and $Q^1(y^0,y^1,v^1) \equiv d^1(y^1,v^1)/d^1(y^0,v^1) = 1/d^1(y^0,v^1) = 0F/0A$. Two more complex cases are $Q^0(y^0,y^1,v^1) \equiv d^0(y^1,v^1)/d^0(y^0,v^1) = [0N/0L]/[0A/0D] = [0N/0L][0D/0A]$ and $Q^1(y^0,y^1,v^0) \equiv d^1(y^1,v^0) = [0N/0J]/[0A/0B] = [0N/0J]/[0B/0A]$.

Are there conditions which will imply the equality of all three classes of output indexes? The three classes of output indexes have their counterparts in consumer theory, and in that context, Pollak [1971; p.65][12] showed that the three classes of consumer quantity indexes coincide if and only if the consumer's preferences were homothetic.[13] The following theorem shows that we have a similar result in the producer context, except that the homotheticity assumption is replaced by a homothetic separability[14] assumption.

Theorem 5: Assume: (i) for $t = 0,1,\ldots,T$ the period t production possibilities set S^t satisfies regularity Conditions I listed in Footnote 6, (ii) $p^0 >> 0_M$ and the period 0 output vector $y^0 > 0_M$ is a solution to the period 0 value added maximization problem $\max_y \{p^0 \cdot y: (y,v^0) \epsilon S^0\} \equiv \pi^0(p^0,v^0)$, and (iii) $p^1 >> 0_M$ and the period 1 output vector $y^1 > 0_M$ is a solution to the period 1 value added maximization problem $\max_y \{p^1 \cdot y: (y,v^1) \epsilon S^1\}$ $\equiv \pi^1(p^1,v^1)$. Then for all reference input vectors $v >> 0_N$, all reference output price vectors $p >> 0_M$ and all time periods $t = 0,1,\ldots,T$, we have

$$Q^t_{FS}(p^0,p^1,v,y^0,y^1) = Q_{SSS}(v^0,v^1,p) = Q^t(y^0,y^1,v) \tag{28}$$

where the Fisher-Shell quantity index Q^t_{FS} is defined by (22), the Samuelson-Swamy-Sato index Q_{SSS} is defined by (23), and the Malmquist-Moorsteen-Bergson index Q^t is defined by (26), if and only if for all $p >> 0_M$, $v >> 0_N$ and $t = 0,1,\ldots,T$, the value added function π^t corresponding to S^t has the following separable functional form:

$$\pi^t(p,v) = b(p)c(t,v). \tag{29}$$

Proof: (29) implies (28). Proving the first equality in (28) is easy:

$$Q^t_{FS}(p^0,p^1,v,y^0,y^1) \equiv [p^1 \cdot y^1/p^0 \cdot y^0]/ [\pi^t(p^1,v)/\pi^t(p^0,v)]$$

$$= [\pi^1(p^1,v^1)/\pi^0(p^0,v^0)][\pi^t(p^0,v)/\pi^t(p^1,v)]$$

using assumptions (ii) and (iii)

$$= c(1,v^1)/c(0,v^0)$$

using (29) and cancelling terms

$$= b(p)c(1,v^1)/b(p)c(0,v^0)$$

$$= \pi^1(p,v^1)/\pi^0(p,v^0) \qquad \text{using (29)}$$

$$\equiv Q_{SSS}(v^0,v^1,p) \qquad \text{using definition (23).}$$

In order to establish the second equality in (28), we must first establish some general relationships between π^t and d^t. Using duality theory,[15] it can be shown that

$$\{y: (y,v)\epsilon S^t\} = \{y: p \cdot y \leq \pi^t(p,v) \text{ for every } p > 0_M\} . \qquad (30)$$

Thus using (25) and (30), for every $v >> 0_N$ and $y >> 0_M$, we have

$$d^t(y,v) \equiv \min_{\delta>0} \{\delta: (y/\delta,v)\epsilon S^t\}$$

$$= \min_{\delta>0} \{\delta: p \cdot y/\delta \leq \pi^t(p,v) \text{ for every } p > 0_M\}$$

$$= \min_{\delta>0} \{\delta: 1/\delta \leq \pi^t(p,v) \text{ for every } p \geq 0_M \text{ such that } p \cdot y = 1\}^{16}$$

$$= 1/\min_p \{\pi^t(p,v): p \cdot y = 1, p \geq 0_M\} . \qquad (31)$$

Using assumption (ii) in the maintained hypotheses of the theorem and (30), we know y^0 solves $\max_y \{p^0 \cdot y: (y,v^0)\epsilon S^0\} = \max_y \{p^0 \cdot y: p \cdot y \leq \pi^0(p,v^0) \text{ for every } p > 0_M\} = \pi^0(p^0,v^0)$. Hence $\pi^0(p,v^0) \geq p \cdot y^0$ for every $p > 0_M$ but $\pi^0(p^0,v^0) = p^0 \cdot y^0$. Thus p^0 solves

$$\min_p \{\pi^0(p,v^0): p \cdot y^0 = p^0 \cdot y^0, p \geq 0_M\} = \pi^0(p^0,v^0),$$

and since $\pi^0(p,v^0)$ is positively linearly homogeneous in p, $p^0/p^0 \cdot y^0$ solves

$$\min_p \left\{\pi^0(p,v^0): p\cdot y^0 = 1,\ p \geq 0_M\right\} = 1 \tag{32}$$

Similarly, using assumption (iii), it can be shown $p^1/p^1\cdot y^1$ solves

$$\min_p \left\{\pi^1(p,v^1):\ \ p\cdot y^1 = 1,\ p \geq 0_M\right\} = 1. \tag{33}$$

Using (29), (32) and (33), it can be seen that $p^t/p^t\cdot y^t$ solves

$$= \min_p\left\{\pi^t(p,v^t): p\cdot y^t = 1,\ p \geq 0_M\right\}$$

$$= \min_p\left\{b(p)c(t,v^t): p\cdot y^t = 1,\ p \geq 0_M\right\}$$

$$= \min_p\left\{b(p): p\cdot y^t = 1,\ p \geq 0_M\right\} c(t,v^t)$$

$$\equiv g(y^t)c(t,v^t) \tag{34}$$

for $t = 0,1$ where the function g is defined by

$$g(y) \equiv \min_p\left\{b(p): p\cdot y = 1,\ p \geq 0_M\right\} \qquad \text{for } y \geq 0_M. \tag{35}$$

Using (31),

$$d^t(y^0,v) = 1/\min_p \left\{\pi^t(p,v): p\cdot y^0 = 1,\ p \geq 0_M\right\}$$

$$= 1/\min_p\left\{b(p): p\cdot y^0 = 1,\ p \geq 0_M\right\}c(t,v) \text{ by (29)}$$

$$= 1/b(p^0/p^0\cdot y^0)c(t,v) \text{ using (34) when } t = 0$$

$$= p^0\cdot y^0/b(p^0)c(t,v) \text{ using the linear homogeneity of b}$$

$$= \pi^0\ (p^0,v^0)/b(p^0)c(t,v) \text{ using assumption (ii)}$$

$$= b(p^0)c(0,v^0)/b(p^0)c(t,v) \text{ using (29).} \tag{36}$$

Using assumption (iii) and (34) when $t = 1$, we may similarly calculate

$$d^t(y^1,v) = b(p^1)c(1,v^1)/b(p^1)c(t,v). \tag{37}$$

By definition (26),

$$Q^t(y^0,y^1,v) \equiv d^t(y^1,v)/d^t(y^0,v)$$

$$= c(1,v^1)/c(0,v^0) \text{ using (36) and (37)}$$

$$= b(p)c(1,v^1)/b(p)c(0,v^0)$$

$$= \pi^1(p,v^1)/\pi^0(p,v^0) \text{ using (29)}$$

$$\equiv Q_{SSS}(v^0,v^1,p) \text{ using definition (33)}$$

which establishes the second equality in (28).

(28) implies (29). Using the first equality in (28), it can be seen that $Q^t_{FS}(p^0,p^1,v,y^0,y^1)$ is independent of v and t. Hence by definitions (2) and (22), for all $t,r = 0,1,\ldots,T$ and all $v >> 0_N$, $\bar{v} >> 0_N$, we have

$$\pi^t(p^1,v)/\pi^t(p^0,v) = \pi^r(p^1,\bar{v})/\pi^r(p^0,\bar{v}). \tag{38}$$

(38) implies that π^t must satisfy (29). Q.E.D.

The above theorem says that all three classes of theoretical output indexes coincide if and only if the technology sets S^t have value added functions π^t that are of the separable

form (29). In the course of proving the above theorem, we found that the deflation or distance function d^t that corresponds to S^t must have the following functional form: for $y >> 0_M$, $v >> 0_N$,

$$d^t(y,v) = 1/g(y)c(t,v) \tag{39}$$

where g was defined by (35). Moreover, $(y,v) >> 0_{M+N}$ will be on the frontier of the period t production possibilities set S^t if and only if $d^t(y,v) = 1$, i.e., if and only if

$$g(y) = 1/c(t,v). \tag{40}$$

A technology set S^t which is such that its efficient output and input combinations (y,x) satisfy an equation of the form (40) is said to have outputs **separable** from inputs.[17] If there is only one input (so N = 1) and no technical progress (so $c(t,v) = c(0,v)$ for all t), then our homothetic separability assumption (29) or (39) reduces to the homotheticity assumption made by Samuelson and Swamy [1974; p.591] when they assert the equality of Q_{SSS} and Q^t in their analysis of index numbers of production possibilities.

Note also that the homothetic separability assumption (29) or (39) implies that $Q^t_{FS}(p^0,p^1,v,y^0,y^1)$ and $Q^t(y^0,y^1,v)$ are independent of t and v and $Q_{SSS}(v^0,v^1,p)$ is independent of p. These results generalize some analogous results due to Sato [1976; pp.438-439] in the context of the Samuelson-Swamy-Sato output index.

Since the homothetic separability assumption (29) or (39) is unlikely to hold empirically, it will be necessary for us to choose one of the three alternative classes of output indexes. Since only the Malmquist output index Q^t satisfies the homogeneity property (27) in general, we shall focus our attention on this index and drop the other two indexes from further consideration.

Theorems 6-10 below provide quantity index counterparts to the price index Theorems 1-4 above.

Theorem 6: Let the reference input vector v and the positive output vectors $y^0 >> 0_M$ and $y^1 >> 0_M$ be given. Suppose $\delta_r \equiv d^t(y^r,v) \equiv \min_{\delta>0} \left\{\delta_0 : (y^r/\delta,v)\epsilon S^t\right\} > 0$ for $r = 0,1$. Suppose S^t satisfies the regularity conditions (i) and (iv) listed in Footnote 6. Then

$$\min_m \left\{y_m^1 / y_m^0 : m=1,\ldots,M\right\} \leq Q^t(y^0,y^1,v) \equiv \delta_1/\delta_0 \leq \max_m \left\{y_m^1 / y_m^0 : \ m=1,\ldots,M\right\}; \tag{41}$$

i.e., the Malmquist output index $Q^t(y^0,y^1,v)$ lies between the smallest and largest ratios of outputs.

Proof: The first inequality in (vi) may be rewritten as

$$y_m^1 / \delta_1 \leq y_m^0 / \delta_0 \qquad \text{for at least one index m.} \tag{42}$$

Suppose (42) were not true. Then we would have

$$y_m^1 / \delta_1 > y_m^0 / \delta_0 \text{ for } m=1,2, \ldots ,M. \tag{43}$$

By the definitions of δ_0 and δ_1, we have $\delta_0 \equiv \min_\delta \left\{\delta: (y^0/\delta,v)\epsilon S^t\right\}$ and $(y^1/\delta_1,v)\epsilon S^t$. But (43) and the free disposal property on S^t (part (iv) of Conditions I) imply that $(y^0/(\delta_0-\epsilon),v)\epsilon S^t$ for some $\epsilon > 0$, which contradicts the definition of δ_0 as a minimum. Thus our supposition is false and the first inequality in (41) follows. The second inequality in (41) follows in an analogous manner, using the minimum nature of δ_1. Q.E.D.

The reader will recall that we singled out two special cases of the class of Fisher-Shell output price indexes: the Laspeyres Fisher-Shell index $P^0(p^0,p^1,v^0)$ defined by (3) and the Paasche Fisher-Shell index $P^1(p^0,p^1,v^1)$ defined by (4). In an analogous manner, we define

the **Laspeyres Malmquist-Bergson-Moorsteen** output index by (44) and the **Paasche Malmquist-Bergson-Moorsteen** output index by (45):

$$Q^0(y^0,y^1,v^0) \equiv d^0(y^1,v^0)/d^0(y^0,v^0) \text{ (period 0 technology and inputs)} \tag{44}$$

$$Q^1(y^0,y^1,v^1) \equiv d^1(y^1,v^1)/d^1(y^0,v^1) \text{ (period 1 technology and inputs).} \tag{45}$$

Theorem 7: Assume: (i) $p^0 >> 0_M$ and $y^0 > 0_M$[18] solves $\max_y \; p^0 \cdot y: (y,v^0)\epsilon S^0$, (ii) $p^1 >> 0_M$ and $y^1 > 0_M$ solves $\max_y \; p^1 \cdot y: (y,v^1)\epsilon S^1$, (iii) $\delta_0 > 0$ solves $\max_{\delta>0} \left\{\delta: (y^1/\delta,v^0)\epsilon S^0\right\} \equiv d^0(y^1,v^0)$ and (iv) $\delta_1 > 0$ solves $\max_{\delta>0} \left\{\delta: (y^0/\delta,v^1)\epsilon S^1\right\} \equiv d^1(y^0,v^1)$. Then

$$Q^0(y^0,y^1,v^0) \geq p^0 \cdot y^1/p^0 \cdot y^0 \equiv Q_L \text{ (the Laspeyres quantity index) and} \tag{46}$$

$$Q^1(y^0,y^1,v^1) \leq p^1 \cdot y^1/p^1 \cdot y^0 \equiv Q_P \text{ (the Paasche quantity index).} \tag{47}$$

Proof: Assumption (i) implies (y^0,v^0) is on the frontier of the set S^0 and hence $d^0(y^0,v^0) \equiv \min_{\delta>0} \left\{\delta: (y^0/\delta,v^0)\epsilon S^0\right\} = 1$. Thus $Q^0(y^0,y^1,v^0) = d^0(y^1,v^0) = \delta_0$ where $(y^1/\delta_0,v^0)\epsilon S^0$ using (iii). Thus y^1/δ_0 is feasible for the maximization problem stated in assumption (i) and so $p^0 \cdot y^0 \geq p^0 \cdot y^1/\delta_0$ or $p^0 \cdot y^1/p^0 \cdot y^0_0 \leq \delta_0 = Q^0(y^0,y^1,v^0)$ which is (46). (47) follows in an analogous manner using assumptions (ii) and (iv) instead of (i) and (iii). Q.E.D.

Corollary 7.1: Q^0 and Q^1 satisfy the following observable bounds if in addition to the hypotheses of the theorem, S^0 and S^1 satisfy (i) and (iv) of Conditions I listed in Footnote 6 and $y^0 >> 0_M$, $y^1 >> 0_M$:

$$Q_L \equiv p^0 \cdot y^1/p^0 \cdot y^0 \leq Q^0(y^0,y^1,v^0) \leq \max_m \left\{y^1_m/y^0_m: \quad m=1,2,\ldots,M\right\}; \tag{48}$$

$$\min_m \left\{y^1_m/y^0_m: \quad m=1,\ldots,M\right\} \leq Q^1(y^1,y^1,v^1) \leq p^1 \cdot y^1/p^1 \cdot y^0 \equiv Q_P. \tag{49}$$

The corollary follows combining Theorems 6 and 7.

The bounds in (48) and (49) will usually be too wide to be of much empirical use. However, Q_L and Q_P will generally be quite close. Hence, it is useful to be able to find a theoretical output index that is bounded by Q_L and Q_P.[19] The dependence of $d^t(y,v)$ on t causes us some problems initially. Hence we first define a deflation function for the economy which is based on a weighted average of the technology sets for periods 0 and 1. Thus for $0 \leq t \leq 1$, $y >> 0_M$ and $v >> 0_N$, define the **"average" technology deflation function** d by

$$d(y,v,t) \equiv \min_{\delta>0}\left\{\delta:\ (y/\delta,v)\varepsilon\ [\ (1-t)S^0 + tS^1]\right\}. \tag{50}$$

Note that the period 0 technology set S^0 gets the weight 1-t in (50) while S^1 gets the weight t, so that $d(y,v,0) = d^0(y,v)$ and $d(y,v,1) = d^1(y,v)$.[20] Given definition (50), we may now define a Malmquist output index for the "average" technology set $(1-t)S^0 + tS^1$ for $0 \leq t \leq 1$ as follows: for positive output vectors $y^0 >> 0_M$, $y^1 >> 0_M$ and a positive reference input vector $v >> 0_N$,

$$Q(y^0,y^1,v,t) \equiv d(y^1,v,t)/d(y^0,v,t). \tag{51}$$

We use the output index (51) in Theorem 8 below.

Theorem 8: Assume that assumptions (i)-(iv) listed in Theorem 7 hold. Suppose in addition that the period 0 technology set S^0 and the period 1 technology set S^1 are sufficiently well-behaved[21] so that the average technology deflation functions $d(y^0,(1-t)v^0+tv^1,t)$ and $d(y^1,(1-t)v^0+tv^1,t)$ are well defined by (50) and continuous positive functions of t for $0 \leq t \leq 1$. Then there exists t^* such that $0 \leq t^* \leq 1$ and the Malmquist output index $Q(y^0,y^1,(1-t^*)v^0 + t^*v^1,t^*)$ defined by (51) for the "average" technology set $(1-t^*)S^0 + t^*S^1$ and using the "average" input vector $(1-t^*)v^0 + t^*v^1$ lies between the Paasche and

Laspeyres output indexes; i.e., if $Q_L \equiv p^0 \cdot y^1 / p^0 \cdot y^0 \le Q_P \equiv p^1 \cdot y^1 / p^1 \cdot y^0$, we have (52), while if $Q_L \ge Q_P$, we have (53):

$$Q_L \le Q(y^0, y^1, (1-t^*)v^0 + t^* v^1, t^*) \le Q_P \text{ or} \tag{52}$$

$$Q_P \le Q(y^0, y^1, (1-t^*)v^0 + t^* v^1, t^*) \le Q_L \tag{53}$$

Proof: Define $h(t) \equiv Q(y^0, y^1, (1-t)v^0 + tv^1, t) \equiv d(y^1, (1-t)v^0 + tv^1, t)/d(y^0, (1-t)v^0 + tv^1, t)$, a continuous function for $0 \le t \le 1$. Note that $h(0) = d(y^1, v^0, 0)/d(y^0, v^0, 0) = d^0(y^1, v^0)/d^0(y^0, v^0) \equiv Q^0(y^0, y^1, v^0) \ge Q_L$ using (46) and $h(1) = d(y^1, v^1, 1)/d(y^1, v^1, 1) = d^1(y^1, v^1)/d^1(y^0, v^1) \equiv Q^1(y^0, y^1, v^1) \le Q_P$ using (47). In view of the inequalities $h(0) \ge Q_L$ and $h(1) \le Q_P$, there are only six possible inequalities between the four numbers $h(0)$, $h(1)$, Q_L and Q_P: see the six cases in Theorem 3 except Q_L replaces P_L and Q_P replaces P_P. The remainder of the proof follows the proof of Theorem 3. Q.E.D.

Theorem 8 is very useful from an empirical point of view since it shows that the observable Paasche and Laspeyres quantity indexes, Q_P and Q_L, (which will usually be quite close to each other) bound a theoretical output index of the form $Q(y^0, y^1, (1-t^*)v^0 + t^* v^1, t^*)$ under rather weak assumptions about the functional forms for the production possibilities sets S^0 and S^1. Hence an average of Q_P and Q_L such as Fisher's [1922] ideal quantity index, $Q_F \equiv (Q_P Q_L)^{1/2}$, should approximate the theoretical index $Q(y^0, y^1, (1-t^*)v^1 + t^* v^1, t^*)$ rather well.

Theorem 9 below requires the definition of the **translog output index** Q_T:

$$Q_T(p^0, p^1, y^0, y^1) \equiv \Pi\, (y_i^1/y_i^0)^{[\frac{1}{2} s_i^0 + \frac{1}{2} s_i^1]} \tag{54}$$

where $s_i^t \equiv p_i^t y_i^t / p^t \cdot y^t$ for $t = 0,1$ and $i = 1,2,\ldots M$.

Theorem 9: (Caves, Christensen and Diewert [1982; p.1401]): Suppose: (i) the deflation functions d^t defined by (25) have the translog functional form for $t = 0$ and 1; i.e., $\ell nd^t(y,v)$ is defined by the right-hand side of (19) for $t = 0$ and 1 except that $y >> 0_M$ replaces p, (ii) the economy or sector faces the positive output price vector $p^t >> 0_M$ and has available the positive input vector $v^t >> 0_N$ during period $t = 0,1$ (iii) $y^t >> 0_M$ solves $\max_y \{p^t \cdot y: (y,v^t) \epsilon S^t\}$ for $t = 0,1$ and (iv) $\alpha^0_{ij} = \alpha^1_{ij}$ for $1 \le i, j \le M$; i.e., the coefficients for the quadratic terms in the logarithms of outputs are identical for the two translog deflation functions d^0 and d^1. Then a geometric average of the theoretical output indexes defined by (44) and (45) is exactly equal to the translog output quantity index defined by (54); i.e.,

$$[Q^0(y^0,y^1,v^0)Q^1(y^0,y^1,v^1)\,]^{1/2} = Q_T\,(p^0,p^1,y^0,y^1). \tag{55}$$

Although Theorem 9 requires a specific functional form for the deflation functions d^0 and d^1, it should be noted that the translog functional form can approximate an arbitrary twice continuously differentiable functional form to the second order.

The comments we made about the numerical approximation properties of the price indexes P_L, P_P, P_F and P_T have their quantity counterparts: Diewert [1978] showed that Q_L, Q_P, Q_T and Q_F will numerically approximate each other to the **first** order when each function is evaluated at a point where $p^0 = p^1$ and $y^0 = y^1$, and P_F and P_T will approximate each other to the **second** order. Thus the theoretical output indexes mentioned in Theorems 8 and 9, $Q(y^0,y^1,(1-t^*)v^0 + t^*v^1,t^*)$ and $[Q^0(y^0,y^1,v^0)Q^1(y^0,y^1,v^1)]^{1/2}$, will both be closely approximated by the observable output indexes, $Q_F(p^0,p^1,y^0,y^1) \equiv [Q_L Q_P]^{1/2}$ and $Q_T(p^0,p^1,y^0,y^1)$. These last two indexes are the output indexes that we would recommend be used in practice.

We conclude this section by reminding the reader that the three classes of output indexes that we have considered in this section, defined by (22), (23) and (51), all have the following property: the contribution to output growth due to the growth of inputs is not distinguished from the contribution due to technical progress. Thus real output may grow a great deal between periods 0 and 1 but this may be due to people working longer hours or due to the growth of inputs in general. For many purposes, we may wish to determine

what portion of the growth in outputs is due to increased efficiency (or technical progress or total factor productivity as it is sometimes called) rather than input growth. This leads us to our next topic, productivity indexes.

4. Productivity Indexes

Recall the definition of the period t (output) deflation function (25): $d^t(y,v) \equiv \min_{\delta>0} \{\delta: (y/\delta,v)\epsilon S^t\}$ where $y >> 0_M$ is an output vector, $v >> 0_N$ is an input vector, and S^t is the technology set for period t.

Let $(y^0,v^0) >> 0_{M+N}$ and $(y^1,v^1) >> 0_{M+N}$ be the observed vectors of outputs and inputs for the economy the periods 0 and 1. Then we may define the period t technology **productivity index** by [22]

$$H^t(x^0,x^1,y^0,y^1) \equiv d^t(y^1,v^1)/d^t(y^0,v^0) \qquad ; t = 0,1 \tag{56}$$

Let us assume that (y^t,v^t) is on the frontier of the period t technology set S^t for $t = 0,1$, so that

$$d^0(y^0,v^0) = 1 \text{ and } d^1(y^1,v^1) = 1. \tag{57}$$

Under these assumptions, the two Hicksian productivity indexes H^t defined by (56) reduce to:

$$\begin{aligned} H^0(v^0,v^1,y^0,y^1) &= d^0(y^1,v^1)/1 \qquad (58) \\ &= \min_{\delta>0}\left\{\delta: \ (y^1/\delta,v^1)\epsilon S^0\right\} \\ &\equiv \delta_0 \qquad \text{say, and} \end{aligned}$$

$$H^1(v^0,v^1,y^0,y^1) = 1/d^1(y^0,v^0) \tag{59}$$

$$= 1/\min_{\delta>0}\left\{\delta:\ (y^0/\delta,v^0)\epsilon S^1\right\}$$

$$\equiv 1/\delta_1 \qquad \text{say.}$$

Thus δ_0 which appears in (58) is the proportional deflation factor which is such that the resulting deflated period 1 output vector y^1/δ_0 would just be producible in period 0 using the period 1 input vector v^1. Thus if $\delta_0 > 1$, there has been technical progress or an increase in efficiency going from period 0 to period 1. Similarly, the δ_1 which appears in (59) is the proportional deflation factor which is such that the resulting deflated period 0 output vector y^0/δ_1 would just be producible in period 1 using the period 0 input vector v^0. Thus if $\delta_1 < 1$, there has been an increase in efficiency going from period 0 to period 1. Thus H^0 and H^1 are alternative measures of efficiency increase. If $H^0 > 1$ or $H^1 > 1$, then we say there has been an increase in efficiency or total factor productivity going from the period 0 technology set S^0 to the period 1 technology set S^1.

As usual, the assumption of maximizing behaviour leads to some interesting (observable) bounds for the theoretical productivity indexes H^0 and H^1 defined by (58) and (59). In this section, the assumption of profit maximizing behaviour replaces the revenue maximizing assumptions that we made in the previous two sections.

Theorem 10: Assume: (i) $p^t >> 0_M$, $w^t >> 0_N$ are given positive vectors of output and input prices for periods $t = 0,1$ and (y^t,v^t) solves the period t profit maximization problem $\max_{y,v}\left\{p^t\cdot y - w^t\cdot v:\ (y,v)\epsilon S^t\right\}$ for $t = 0,1$ (and this in turn implies that (57) will hold), (ii) δ_0 solves the minimization problem in (58), (iii) $\delta_1 > 0$ solves the minimization problem in (59), and (iv) $p^1\cdot y^0 > 0$ and $p^0\cdot y^1 > 0$. Then the productivity indexes H^0 and H^1 defined by (58) and (59) satisfy the following bounds:

$$H^1(v^0,v^1,y^0,y^1) \le \frac{p^1\cdot y^1}{p^1\cdot y^0} + \frac{w^1\cdot v^1}{p^1\cdot y^0}\left[\left(\frac{w^1\cdot v^1}{w^1\cdot v^0}\right)^{-1} - 1\right] \equiv B_1; \tag{60}$$

$$H^0(v^0,v^1,y^0,y^1) \geq \left\{ \left(\frac{p^0 \cdot y^1}{p^0 \cdot y^0}\right)^{-1} + \frac{w^0 \cdot v^0}{p^0 \cdot y^1}\left[\left(\frac{w^0 \cdot v^1}{w^0 \cdot v^0}\right) - 1 \right] \right\}^{-1} \equiv B_0; \tag{61}$$

Proof: By (iii), $(y^0/\delta_1, v^0) \epsilon S^1$. Hence since by (i) $(y^1,v^1) \epsilon S^1$ solves the period 1 profit maximization problem, we must have

$$p^1 \cdot y^1 - w^1 \cdot v^1 \geq p^1 \cdot (y^0/\delta_1) - w^1 \cdot v^0. \tag{62}$$

Using (59) and (iv), $p^1 \cdot y^0 > 0$, (62) may be rearranged to yield (60). (61) follows in an analogous manner, using (ii), (i), (58) and $p^0 \cdot y^1 > 0$. Q.E.D.

Note that the Paasche index for outputs $p^1 \cdot y^1/p^1 \cdot y^0$ and the Paasche index for inputs $w^1 \cdot v^1/w^1 \cdot v^0$ appear in (60) while the Laspeyres output and input indexes, $p^0 \cdot y^1/p^0 \cdot y^0$ and $w^0 \cdot v^1/w^0 \cdot v^0$ respectively, appear in (61). If $w^1 \cdot v^1/p^1 \cdot y^0 > 0$, $p^1 \cdot y^1/p^1 \cdot y^0 > 1$ and $w^1 \cdot v^1/w^1 \cdot v^0 \leq 1$, then $B_1 > 1$ while if $w^0 \cdot v^0/p^0 \cdot y^1 > 0$, $p^0 \cdot y^1/p^0 \cdot y^0 > 1$ and $w^0 \cdot v^1/w^0 \cdot v^0 \leq 1$, then $B_0 > 1$. Thus if the Paasche and Laspeyres output indexes exceed unity while the Paasche and Laspeyres input indexes are less than unity, then both productivity bounds B_0 and B_1 will exceed unity.

If there are constant returns to scale in production, then profits will be zero in the two periods and so $p^1 \cdot y^1 = w^1 \cdot v^1$ and $p^0 \cdot y^0 = w^0 \cdot v^0$. Under these conditions, we find that $B_1 = (p^1 \cdot y^1/p^1 \cdot y^0)/(w^1 \cdot v^1/w^1 \cdot v^0)$ = Paasche output index divided by the Paasche input index and $B_0 = (p^0 \cdot y^1/p^0 \cdot y^0)/(w^0 \cdot v^1/w^0 \cdot v^0)$ = Laspeyres output index divided by the Laspeyres input index.

We can of course invent a theoretical productivity index that will lie between the numbers B_1 and B_0 that are defined in (60) and (61). We need only recall the definition (50) of the "average" technology deflation function $d(y,v,t)$(which gives the period 0 technology set the weight (1-t) and the period 1 technology set the weight t) and define the **Hicks productivity index** for the "average" technology set $(1\text{-}t)S^0 + tS^1$ as

$$H(v^0,v^1,y^0,y^1,t) \equiv d(y^1,v^1,t)/d(y^1,v^0,t). \tag{63}$$

Note that $H(v^0,v^1,y^0,y^1,0) = H^0(v^0,v^1,y^0,y^1)$ and $H(v^0,v^1,y^0,y^1,1) = H^1(v^0,v^1,y^0,y^1)$ where H^0 and H^1 are defined by (56). In general, $H(v^0,v^1,y^0,y^1,t)$ is a productivity index like $H^t(v^0,v^1,y^0,y^1)$ defined by (56) except $(1-t)S^0 + tS^1$ is the reference technology set in (63), while S^t is the reference technology set in (56).

Theorem 11: Assume the regularity conditions made in Theorem 10 and the following additional assumptions: (i) $y^t >> 0_M$, $v^t >> O_N$ for $t = 0,1$ and (ii) $d(y^1,v^1,t)$ and $d(y^0,v^0,t)$ are positive, continuous functions of t for $0 \le t \le 1$.[23] Then there exists t^* such that $0 \le t^* \le 1$ and the theoretical productivity index $H(v^0,v^1,y^0,y^1,t^*)$ defined by (63) lies between the observable bounds B_1 and B_0 defined in (60) and (61).

Proof: Define $h(t) \equiv H(v^0,v^1,y^0,y^1,t) \equiv d(y^1,v^1,t)/d(y^0,v^0,t)$, a continuous function for $0 \le t \le 1$. Note that $h(0) \equiv H(v^0,v^1,y^0,y^1,0) = H^0(v^0,v^1,y^0,y^1) \ge B_0$ using (61) and $h(1) \equiv H(v^0,v^1,y^0,y^1) = H^1(v^0,v^1,y^0,y^1) \le B_1$ using (60). The remainder of the proof follows the proof of Theorem 3 except that B_0 replaces P_L and B_1 replaces P_P.Q.E.D.

Theorem 11 is useful only if the bounds B_0 and B_1 are close to each other. How close will they be in practice? Regard B_0 and B_1 as functions of all of the price and quantity vectors for the two periods; i.e., define $B_1(v^0,v^1,y^0,y^1,w^0,w^1,p^0,p^1)$ by the right-hand side of (60), and define $B_0(v^0,v^1,y^0,y^1,w^0,w^1,p^0,p^1)$ by the right-hand side of (61). Then straightforward but tedious calculations show that the level and all of the first order partial derivatives of the functions B_0 and B_1 coincide when evaluated at an equal price and quantity point; i.e., at a point where $v^0 = v^1$, $y^0 = y^1$, $w^0 = w^1$ and $p^0 = p^1$. Thus B_0 and B_1 numerically approximate each other to the first order when evaluated at an equal price and quantity point, and thus they should normally be about as close to each other as the Paasche and Laspeyres quantity indexes. Hence, taking an average of B_0 and B_1, such as $(B_0B_1)^{1/2}$ or $\frac{1}{2}B_0 + \frac{1}{2}B_1$, should yield an adequate approximation to the theoretical productivity index $H(x^0,x^1,y^0,y^1,t^*)$.

Theorem 11 was established without making any specific assumptions about the functional form for the period 0 and 1 deflation functions $d^0(y,v)$ and $d^1(y,v)$. In the following theorem, we assume d^0 and d^1 have the translog functional form. As usual, we cannot identify the two individual Hicksian productivity indexes H^0 and H^1 defined by (58) and

(59), but we can identify their geometric average, $(H^0H^1)^{1/2}$.

Theorem 12: (Caves, Christensen and Diewert [1982; pp.1404-1406]): Assume: (i) the technology sets S^0 and S^1 are such that the output deflation functions $d^0(y,v)$ and $d^1(y,v)$ defined by (25) have the translog functional form; i.e., $\ell n d^t(y,v)$ is defined by the right-hand side of (19) for t = 0 and 1 except that $y >> 0_M$ replaces $p >> 0_M$, (ii) $p^t >> 0_M$, $w^t >> 0_N$ are given positive vectors of output and input prices for periods t = 0,1 and $(y^t,v^t) > 0_{M+N}$ solves the period t profit maximization problem, $\max_{y,v} \{p^t \cdot y - w^t \cdot v: (y,v) \epsilon S^t\}$ for t = 0,1, and (iii) the coefficients pertaining to the quadratic terms of the two translog deflation functions are identical; i.e.,

$$\alpha^0_{ij} = \alpha^1_{ij},\ 1 \leq i, j \leq M;\ \beta^0_{ij} = \beta^1_{ij},\ 1 \leq i, j \leq N; \tag{64}$$

$$\gamma^0_{mn} = \gamma^1_{mn},\ m = 1,\ldots,M,\ n = 1,\ldots,N.$$

Then the geometric mean of the two theoretical productivity indexes H^0 and H^1 defined by (58) and (59) may be calculated by the right-hand side of (65):

$$[H^0(v^0,v^1,y^0,y^1)H^1(v^0,v^1,y^0,y^1)]^{1/2} \tag{65}$$

$$= \frac{\Pi^M_{m=1}(y^1_m/y^0_m)^{\frac{1}{2}[(p^0_m y^0_m/p^0\cdot y^0) + (p^1_m y^1_m/p^1\cdot y^1)]}}{\Pi^N_{n=1}(v^1_n/v^0_n)^{\frac{1}{2}[(w^0_n v^0_n/p^0\cdot y^0) + (w^1_n v^1_n/p^1\cdot y^1)]}} \equiv H_T$$

Corollary 12.1: (Caves, Christensen and Diewert [1982; p.1406]): If profits are zero in the two periods so that $p^t \cdot y^t = w^t \cdot v^t$ for t = 0,1 (this is consistent with there being constant returns to scale in production), then the right-hand side of (65) becomes $Q_T(p^0,p^1,y^0,y^1)/Q_T(w^0,w^1,v^0,v^1)$, the ratio of the translog output index to the translog input index.

The proofs of the above theorem and corollary are rather technical and hence will not be reproduced here.[24] Note that the translog output index was defined by (54) and the

translog input index is defined in an analogous manner. It should also be noted that although the quadratic coefficients of the two translog deflation functions are restricted to be equal (recall (64)), the linear terms are unrestricted. Hence the functional form assumptions made in Theorem 12 are not that restrictive from an empirical point of view.

Let us define the observable productivity index on the right-hand side of (65) as $H_T(p^0,p^1,y^0,y^1,w^0,w^1,v^0,v^1)$, the **translog productivity index**. Define the **nonparametric productivity index** H_F[25] as the geometric mean of the productivity index bounding functions $B_1(p^1,y^0,y^1,w^1,v^0,v^1)$ and $B_0(p^0,y^0,y^1,w^0,v^0,v^1)$ defined in (60) and (61):

$$H_F(p^0,p^1,y^0,y^1,w^0,w^1,v^0,v^1) \equiv [B_0B_1]^{1/2}. \qquad (66)$$

Theorem 11 would lead us to calculate H_F as an empirically computable productivity index for the economy going from period 0 to period 1, while Theorem 12 would lead us to use H_T as an appropriate productivity index for the economy. How close will H_T be to H_F? The following theorem indicates that they will be the same for all practical purposes.

Theorem 13: H_T and H_F differentially approximate each other to the second order at any point where the prices and quantities are pertaining to one period equal the prices and quantities pertaining to the other period; i.e., we have $H_F(p^0,p^1,y^0,y^1,w^0,w^1,v^0,v^1) = H_T(p^0,p^1,y^0,y^1,w^0,w^1,v^0,v^1)$, $\partial H_F(p^0,p^1,y^0,y^1,w^0,w^1,v^0,v^1)/\partial z_i = \partial H_T(p^0,p^1,y^0,y^1,w^0,w^1,v^0,v^1)/\partial z_i$ and $\partial^2 H_F/\partial z_i\partial z_j = \partial^2 H_T/\partial z_i\partial z_i$ where z_i and z_j are any components of the vector $(p^0,p^1,y^0,y^1,w^0,w^1,v^0,v^1)$ provided that all of the derivatives are evaluated at a point where $p^0 = p^1 >> 0_M$, $y^0 = y^1 >> 0_M$, $w^0 = w^1 >> 0_N$ and $v^0 = v^1 >> 0_N$.

The proof of the above theorem is a series of straightforward but tedious computations. In the zero profits case where $p^t \cdot y^t = w^t \cdot v^t$ for $t = 0,1$, Theorem 13 may be proven using Theorem 6 in Diewert [1978; p.888].

In this section and in the previous two sections, we have defined certain theoretical classes of: (i) output price indexes, (ii) output quantity indexes and (iii) productivity indexes. In

each case, we have applied two approaches in order to obtain empirically computable index number formulae that will approximate an appropriate theoretical index. These two approaches might be called: (1) the nonparametric approach that relies on various theoretical bounds and (2) the translog approach which assumes that certain functional forms are translog. We found that in each case, our two approaches led to very similar index number formulae, which approximated each other to the second order in a certain sense. We recommend that either of the two index number formulae be used in empirical situations.

There is at least one major problem that we have ignored up to now: the existence of internationally traded goods. We consider this omission in the following section.

5. The Adjustment of Output and Productivity Indexes for Changes in Terms of Trade

Much of international trade theory assumes that internationally traded goods may be directly bought and sold by both producers and consumers in the home country. We will not take this point of view; instead we shall assume that there are J additional goods in the economy that are exported and K additional goods that are imported. Thus we assume that all imported goods that arrive at the border of the home country have additional domestic transportation, wholesaling, retailing or processing inputs added to them before these imported goods are sold to consumers, while exported goods have different transportation requirements than competing domestic goods. For an empirical application of this treatment of traded goods, see Kohli [1978].

In Section 3, we derived indexes to measure an economy's growth of real output. In Section 4, we derived a class of productivity indexes that essentially allowed us to decompose the growth of outputs into two terms: one that reflected the growth in inputs[26] and one that reflected the growth in efficiency (a productivity index). In this section, we wish to find the contribution to an economy's output growth rate between two periods due to: (i) input growth, (ii) improvements in efficiency, (iii) changes in export prices, (iv) changes in import prices and (v) changes in the economy's (private production sector) balance of trade deficit. The last three sources of output growth will sometimes be grouped into one source in what follows: changes in the economy's terms of trade. There is probably no

need to stress the practical importance of having accurate methods of measuring the contribution to real output growth of changes in an open economy's terms of trade: governments, unions and businesses are all interested in this measurement problem.

We can now explain why we choose to treat internationally traded goods as separate goods which interact directly only with the production sector of the economy, rather than with both the production and consumer sectors: our production theory oriented treatment of traded goods leads to a fairly simple method for attributing a portion of the economy's output growth to changes in the international environment, using only producer theory. The traditional treatment of internationally traded goods, which assumes each traded good is directly available to both producers and consumers, requires a complete general equilibrium model in order to measure the effects of changes in the terms of trade. Our producer oriented approach not only requires less information to implement, but it is probably more realistic as well.

In the previous sections, we assumed that the economy had a period t technology set S^t for periods $t = 0$ and 1. We now assume that the private production sector of the economy has the technology set T^0 in period 0 and T^1 in period 1. Points belonging to T^t are of the form (y,v,x,m) where $y \geq 0_M$ is an n dimensional vector of domestic **outputs**, $v \geq 0_N$ is an N dimensional vector of domestic (variable) **inputs**, $x \geq 0_J$ is a J dimensional vector of **exports**, and $m \geq 0_K$ is a K dimensional vector of **imports**. Throughout this section, we shall assume that the technology sets T^t satisfy Conditions I on the sets S^t.[27]

When we defined the national product function π^t by (1), we maximized the value of output that could be produced by the period t technology set S^t at output prices p, conditional on the vector of variable inputs v being available. In (67) below, we define the economy's **value added function** ϕ^t in a similar manner, except that we condition on the export vector x and the import vector m as well as the input vector v; i.e., for any vector of positive output prices $p >> 0_M$, nonnegative input vector $v \geq 0_N$, nonnegative export vector $x \geq 0_J$, nonnegative import vector $m \geq 0_K$ and technology set indexed by t, define[28]

$$\phi^t(p,v,x,m) \equiv \max_y \left\{p \cdot y: \quad (y,v,x,m) \epsilon T^t\right\}, \; t = 0,1. \tag{67}$$

Given that the technology set T^t satisfies Conditions IV, the value added function ϕ^t will satisfy Conditions V, which are similar to Conditions II on π^t.[29]

At this point, we could rework our material on output price indexes, output quantity indexes and productivity indexes presented earlier in Sections 2, 3 and 4 with ϕ^t replacing π^t. Fortunately, there is no need to do this in detail: all we need to do is to define S^t in terms of T^t by (68) and π^t in terms of ϕ^t by (69) below, and then we simply reinterpret our old indexes in the light of these new definitions:

$$S^t \equiv \left\{(y,v): \quad (y,v,x^t,m^t) \epsilon T^t\right\}, \; t = 0,1; \tag{68}$$

$$\pi^t \,(p,v) \equiv \phi^t(p,v,x^t,v^t), \; t = 0,1 \tag{69}$$

where $x^t \geq 0_J$ and $m^t \geq 0_K$ are observed export and import vectors for the economy during periods $t = 0,1$. Thus the period t **restricted production possibilities** S^t defined by (68) is simply the set of domestic output and input combinations that can be produced by the (complete) period t technology set T^t if we restrict the economy to produce the observed period t export vector x^t and to utilize the observed period t vector m^t. Thus the transformation of S^0 into S^1 reflects not only changes in the efficiency of the economy (technical progress), but also changes in the international trade environment that is facing the economy during the two periods. Hence if we reinterpret the results of Sections 2-4 using definitions (68) and (69), it can be seen that our indexes are conditional on the vectors of exports and imports that the economy produced and utilized during the two periods. This is not the most natural way of proceeding. For a small open economy, it may be more useful to condition on the structure of foreign prices facing the economy rather than on the export-import vectors. These considerations lead us to define a new value added function, one that holds foreign prices rather than quantities constant.

Let $p >> 0_M$ denote a positive vector of domestic output prices and let $v \geq 0_N$ denote a nonnegative vector of domestic inputs that are available for use. Let $\rho \equiv (\rho_1,\rho_2,\ldots,\rho_J)$

$>> 0_J$ denote a vector of positive prices for export goods that the private sector of the economy faces, let $\omega \equiv (\omega_1,\omega_2,\ldots,\omega_K)$ denote a vector of positive prices for imported goods that private producers must pay[30] and let the scalar b denote the balance of trade deficit that the private sector of the economy is allowed to accumulate during the period.[31] Then for $t = 0,1$, define the period t **balance of trade restricted value added function** g^t by[32]

$$g^t(p,v,\rho,\omega,b) \equiv \max_{y,x,m}\left\{p\cdot y: \; (y,v,x,m)\epsilon T^t; \rho\cdot x - \omega\cdot m + b \geq 0\right\}. \quad 0 . \tag{70}$$

Thus $g^t(p,v,\rho,\omega,b)$ tells us how much domestic output the economy can produce (valued at the reference prices p) using the period t technology set and the reference vector of domestic inputs v, given that exports may be sold at prices ρ, imports may be purchased at prices ω and the economy is allowed to earn a balance of trade deficit of size b. This function is the producer theory counterpart to Woodland's [1980] (general equilibrium) indirect trade utility function. If T^t satisfies Conditions V, then g^t will satisfy various regularity conditions which are relegated to a footnote.[33]

Let us assume that the economy's observed price and quantity vectors for periods 0 and 1 satisfy the following restrictions:

$$p^t >> 0_M,\ y^t > 0_M,\ v^t >> 0_N,\ \rho^t >> 0_J,\ x^t > 0_J,\ \omega^t >> 0_K,\ m^t > 0_K;\ t = 0,1. \tag{71}$$

We may again reinterpret the indexes developed in Sections 2-4 by defining $\tilde{S}^t$ in terms of T^t (and the structure of foreign prices facing the economy during period t) by (72) below and by defining $\tilde{\pi}^T$ in terms of $\tilde{g}^t$ by (73) below:

$$\tilde{S}^t \equiv \left\{(y,v): \; (y,v,x,m)\epsilon T^t;\ p^t\cdot x - \omega^t\cdot m + b^t \geq 0\right\},\ t = 0,1; \tag{72}$$

$$\tilde{\pi}^t(p,v) \equiv g^t(p,v,p^t,\omega^t,b^t),\ t = 0,1, \tag{73}$$

where b^t is the economy's actual balance of trade deficit during period t; i.e.,

$$b^t = -\rho^t\cdot x^t + \omega^t\cdot m^t,\ t = 0,1. \tag{74}$$

Now let $\tilde{S}^t$ replace S^t and $\tilde{\pi}^t$ in Sections 2-4, and all of our old analysis may be interpreted. Thus we obtain a "new" output price index concept $\tilde{P}^t(p^0,p^1,v)$ defined by (2) except $\tilde{\pi}^t$ replaces π^t, and "new" quantity index concepts $\tilde{Q}^t_{FS}(p^0,p^1,v,y^0,y^1) \equiv p^1 \cdot y^1/p^0 \cdot y^0 \tilde{P}^t(p^0,p^1,v)$ (recall (22)), $\tilde{Q}_{SSS}(v^0,v^1,p) \equiv \tilde{\pi}^1(p,v^1)/\tilde{\pi}^0(p,v^0)$ (recall (23)) and $\tilde{Q}^t(y^0,y^1,v) \equiv \tilde{d}^t(y^1,v)/\tilde{d}^t(y^0,v)$ (recall (26)) where $\tilde{d}^t(y,v) \equiv \min_{\delta>0} \left\{\delta: (y/\delta,v)\epsilon\tilde{S}^t\right\}$. The "new" output price index concept $\tilde{P}^t(p^0,p^1,v)$ is essentially due to Archibald [1977; pp.60-61], who called it the fixed cost output price index.

Recall that we obtained various bounds and approximations to the classes of theoretical indexes that were studied in Sections 2-4. It is interesting to note that we obtain **precisely** the same bounds and approximations to the theoretical indexes no matter whether we use the set S^t and function π^t defined by (68) and (69) or the $\tilde{S}^t$ and function $\tilde{\pi}^t$ defined by (72) and (73). However, it is possible to show that there are certain bounding relationships between the two classes of theoretical indexes. We briefly outline these bounds.

Assume (74) holds. Define $\tilde{S}^t$ by (72) and S^t by (68). Then it is easy to see that:

$$S^t \text{ is a subset of } \tilde{S}^t \text{ for } t = 0,1. \tag{75}$$

Using (67), (68), (69), (70), (72), (73) and (75), we find that since $\tilde{S}^t$ is bigger than S^t,

$$\tilde{\pi}^t(p,v) \geq \pi^t(p,v) \text{ for } t = 0,1.^{34} \tag{76}$$

Using (75) and the definitions of d^t (see (20)) and $\tilde{d}^t$, we have

$$\tilde{d}^t(y,v) \leq d^t(y,v) \text{ for } t = 0,1. \tag{77}$$

In addition to (71) and (74), assume that (y^t,x^t,m^t) solves

$$\max_{y,x,m}\left\{ p^t\cdot y:\ (y,v^t,x,m)\epsilon T^t;\ \rho^t\cdot x - \omega^t\cdot m + b^t \geq 0\right\} = p^t\cdot y^t \tag{78}$$

$$\equiv g^t(p^t,v^t\rho^t,\omega^t,b^t)$$

$$\equiv \tilde{\pi}^t(p^t,v^t)$$

$$= \phi^t(p^t,v^t,x^t,m^t)$$

$$\equiv \pi^t(p^t,v^t) \qquad , t = 0,1.$$

Assumption (78) also implies

$$1 = d^t(y^t,v^t) = \tilde{d}^t(y^t,v^t) \qquad , t = 0,1. \tag{79}$$

Under the above hypotheses, Theorem 2 may be extended to yield the following inequalities due to Archibald [1977; pp.66]:

$$p^1\cdot y^0/p^0\cdot y^0 \leq P^0(p^0,p^1,v^0) \equiv \pi^0(p^1,v^0)/\pi^0(p^0,v^0) \tag{80}$$

$$\leq \tilde{\pi}^0(p^1,v^0)/\tilde{\pi}^0(p^0,v^0) \qquad \text{using (76) and (78)}$$

$$\equiv \tilde{P}^0(p^0,p^1,v^0);$$

$$\tilde{P}^1(p^0,p^1,v^1) \equiv \tilde{\pi}^1(p^1,v^1)/\tilde{\pi}^1(p^0,v^1) \tag{81}$$

$$\leq \pi^1(p^1,v^1)/\pi^1(p^0,v^1) \qquad \text{using (76) and (78)}$$

$$\equiv P^1(p^0,p^1,v^1)$$

$$\leq p^1\cdot y^1/p^0\cdot y^1.$$

In a similar manner, we deduce that the Samuelson-Swamy-Sato output indexes satisfy the following inequalities:

$$p^0 \cdot y^1/p^0 \cdot y^0 \le \pi^1(p^0,v^1)/\pi^0(p^0,v^0) \quad \text{using (78) and the feasibility of } y^1 \text{ for the maximization problem defining } \pi^1(p^0,v^1) \tag{82}$$

$$\equiv Q_{SSS}(v^0,v^1,p^0)$$

$$\le \tilde{\pi}^1(p^0,v^1)/\tilde{\pi}^0(p^0,v^0) \quad \text{using (76) and (78)}$$

$$\equiv \tilde{Q}_{SSS}(v^0,v^1,p^0);$$

$$p^1 \cdot y^1/p^1 \cdot y^0 \ge \pi^1(p^1,v^1)/\pi^0(p^1,v^0) \quad \text{using (78) and the feasibility of } y^0 \text{ for the maximization problem defining } \pi^1(p^1,v^0) \tag{83}$$

$$\equiv Q_{SSS}(v^0,v^1,p^1)$$

$$\ge \tilde{\pi}^1(p^1,v^1)/\tilde{\pi}^0(p^1,v^0) \quad \text{using (78) and (76)}$$

$$\equiv \tilde{Q}_{SSS}(v^0,v^1,p^1);$$

Finally, we may rework Theorem 7 to obtain the following bounds:

$$p^0 \cdot y^1/p^1 \cdot y^0 \le \tilde{Q}^0(y^0,y^1,v^0) \quad \text{using Theorem 7} \tag{84}$$

$$\equiv \tilde{d}^0(y^1,v^0)/\tilde{d}^0(y^0,v^0)$$

$$\le d^0(y^1,v^0)/d^0(y^0,v^0) \quad \text{using (77) and (79)}$$

$$\equiv Q^0(y^0,y^1,v^0);$$

$$p^1 \cdot y^1 / p^1 \cdot y^0 \geq \tilde{Q}^1(y^0,y^1,v^1) \qquad \text{using Theorem 7} \qquad (85)$$

$$\equiv \tilde{d}^1(y^1,v^1)/\tilde{d}^1(y^0,v^1)$$

$$\geq d^1(y^1,v^1)/d^1(y^0,v^1) \qquad \text{using (77) and (79)}$$

$$\equiv Q^1(y^0,y^1,v^1).$$

Our conclusion at this point is that the indexes developed previously in Sections 2-4 above may be reinterpreted in two ways: one way using the S^t restricted production possibilities sets defined by (68) and another way using the $\tilde{S}^t$ sets defined by (72). However, the bounds and empirical approximations for each of the two classes of theoretical indexes turn out to be identical, except that we obtain the bounds (80)-(85) in addition to our previous bounds.

We are finally ready to study our problem of primary concern: how can we determine what portion of the growth in output between the two periods is due to changes in the terms of trade?

Recall the definition of the period t technology balance of trade restricted value added function g^t defined above by (70). We shall use the function g^t in order to define the period t technology **real output growth adjustment factor** A^t due to changes in terms of trade: for each positive reference output price vector $p >> 0_M$, positive reference input vector $v >> 0_N$, period 0 and 1 positive export price vectors $p^0 >> 0_J$ and $p^1 >> 0_J$, period 0 and 1 positive import price vectors $\omega^0 >> 0_K$, and $\omega^1 >> 0_K$, period 0 and 1 allowable balance of trade deficits b^0 and b^1 and for $t = 0$ or 1, define the terms of trade adjustment factor A^t by

$$A^t(\rho^0,\omega^0,b^0,\rho^1,\omega^1,b^1,p,v) \equiv g^t(p,v,\rho^1,\omega^1,b^1)/g^t(p,v,\rho^0,\omega^0,b^0). \qquad (86)$$

Thus A^t defined by (86) represents the percentage increase in the economy's real domestic output (valued at the reference output prices p) that can be attributed to a change in export and import prices from (ρ^0, ω^0) to (ρ^1, ω^1) and a change in the economy's balance

of trade deficit from b^0 to b^1, holding domestic input utilization constant at the reference input vector v and utilizing the economy's period t technology set T^t.

As is usual in index number theory, two special cases of the family of theoretical indexes A^t defined by (86) will be extremely useful, a Laspeyres and a Paasche type theoretical adjustment factor:

$$\bar{A}^0 \equiv A^0(\rho^0,\omega^0,b^0,\rho^1,\omega^1,b^1,p^0,v^0) \text{ and } \bar{A}^1 \equiv A^1(\rho^0,\omega^0,b^0,\rho^1,\omega^1,b^1,p^1,v^1) \quad (87)$$

Let us assume (71), (74) and (78). Then the ratio of the value of output produced by the economy in the two periods may be written as follows:

$$p^1 \cdot y^1/p^0 \cdot y^0 = g^1(p^1,v^1,\rho^1,\omega^1,b^1)/g^0(p^0,v^0,\rho^0,\omega^0,b^0)$$

$$= \frac{g^1(p^1,v^1,\rho^1,\omega^1,b^1)}{g^1(p^1,v^1,\rho^0,\omega^0,b^0)} \frac{g^1(p^1,v^1,\rho^0,\omega^0,b^0)}{g^0(p^1,v^0,\rho^0,\omega^0,b^0)} \frac{g^0(p^1,v^0,\rho^0,\omega^0,b^0)}{g^0(p^0,v^0,\rho^0,\omega^0,b^0)}$$

$$\equiv \bar{A}^1 \tilde{Q}_S(v^0,v^1,p^1) \tilde{P}^0(p^0,p^1,v^0) \quad (88)$$

where $\bar{A}^1$ is the Paasche terms of trade adjustment factor defined by (87), $\tilde{Q}_S(v^0,v^1,p^1)$ is a variant of the Paasche Samuelson-Swamy-Sato output index (23) (the variant holds the international environment facing the economy constant) and $\tilde{P}^0(p^0,p^1,v^0)$ is the Laspeyres Fisher-Shell output price index (3) that corresponds to $\tilde{\pi}^t$ defined by (73). Thus the actual nominal private value of domestic production ratio for the two periods, $p^1 \cdot y^1/p^0 \cdot y^0$, may be written as the product of a terms of trade adjustment factor times an output quantity index times an output price index.

The nominal output ratio may be decomposed in an alternative manner as follows:

$$p^1 \cdot y^1 / p^0 \cdot y^0 = g^1(p^1,v^1,\rho^1,\omega^1,b^1)/g^0(p^0,v^0,\rho^0,\omega^0,b^0)$$

$$= \frac{g^0(p^0,v^0,\rho^1,\omega^1,b^1)}{g^0(p^0,v^0,\rho^0,\omega^0,b^0)} \; \frac{g^1(p^0,v^1,\rho^1,\omega^1,b^1)}{g^0(p^0,v^0,\rho^1,\omega^1,b^1)} \; \frac{g^1(p^1,v^1,\rho^1,\omega^1,b^1)}{g^1(p^0,v^1,\rho^1,\omega^1,b^1)}$$

$$\equiv \bar{A}^0 \, \tilde{Q}_S(v^0,v^1,p^0) \, \tilde{P}^1(p^0,p^1,v^1) \tag{89}$$

where $\bar{A}^0$ is the Laspeyres terms of trade adjustment factor defined in (87), $\tilde{Q}_S(v^0,v^1,p^0)$ is a variant of the Laspeyres Samuelson-Swamy-Sato output index and $\tilde{P}^1(p^0,p^1,v^1)$ is the Paasche Fisher-Shell output price index.

Finally, we may multiply the right-hand sides of (88) and (89) and take the square root of the product to obtain a third decomposition of the nominal output ratio:

$$p^1 \cdot y^1 / p^0 \cdot y^0 = [\bar{A}^1 \bar{A}^0]^{1/2} \, [\tilde{Q}_S(v^0,v^1,p^1) \tilde{Q}_S(v^0,v^1,p^0)]^{1/2} [\tilde{P}^0 \tilde{P}^1]^{1/2}. \tag{90}$$

We shall attempt to find empirically implementable approximations to the theoretical terms of trade adjustment indexes $\bar{A}^1$, $\bar{A}^0$ and $(\bar{A}^1\bar{A}^0)^{1/2}$ that appear in equations (88)-(90) above. Once we have these approximations and approximations to the price indexes appearing in (88)-(90), the quantity indexes which appear in (88)-(90) may be approximated residually. Recall that Theorem 4 provided a good approximation to $[\tilde{P}^0\tilde{P}^1]^{1/2}$.

In order to find useful approximations for the theoretical indexes $\bar{A}^1$ and $\bar{A}^0$, it is necessary to be able to compute the first order partial derivatives of the balance of trade restricted value added functions $g^t(p,v,\rho,\omega,b)$ with respect to the international variables p, ω and b. The following theorem is useful in this respect.

Theorem 14: Assume (i) (71), (74) and (78) hold; the value added functions $\phi^t(p,v,x,m)$ defined by (67) for $t = 0,1$ have the following properties: (ii) $\phi^t(p^t,v^t,x,m)$ is twice continuously differentiable with respect to the components of x and m in a neighbourhood

of x^t, m^t for $t = 0,1$ and (iii) the matrix of second order partial derivatives of $\phi^t(p^t,v^t,x^t,m^t)$ with respect to the components of $z \equiv (x,m)$ is negative definite in the subspace orthogonal to the vector of first order partial derivatives of $\phi^t(p^t,v^t,x^t,m^t)$ with respect to the components of $z \equiv (x,m)$ for $t = 0,1$, (iv) $x^t >> 0_J$, $m^t >> 0_K$ for $t = 0,1$ and (v) $\nabla_x \phi^t(p^t,v^t,x^t,m^t) > 0_J$ for $t = 0,1$. Then: (a) x^t, m^t solves the constrained maximization problem

$$\max_{x,m}\left\{\phi^t(p^t,v^t,x,m):\ \rho^t \cdot x - \omega^t \cdot m + b^t \geq 0\right\} = p^t \cdot y^t \qquad (91)$$

for $t = 0,1$; (b) the Lagrange multiplier for (91) is $e^t > 0$ for $t = 0,1$ and (c) the first order partial derivatives of g^t with respect to the international trade variables are given by:

$$\nabla_\rho g^t(p^t,v^t,\rho^t,\omega^t,b^t) = e^t x^t; \qquad (92)$$

$$\nabla_\omega g^t(p^t,v^t,\rho^t,\omega^t,b^t) = -e^t m^t; \qquad (93)$$

$$\nabla_b g^t(p^t,v^t,\rho^t,\omega^t,b^t) = e^t; \qquad t = 0,1. \qquad (94)$$

Proof: It is clear that the constrained maximization problems in (78) may be rewritten as the constrained maximization problems in (91). Letting e^t be the optimal Lagrange multiplier for the tth problem in (91), our differentiability assumptions on ϕ^t and the positivity assumptions (iv) imply that the following first order conditions for (91) will be satisfied:

$$\nabla_x \phi^t(p^t,v^t,x^t,m^t) + e^t \rho^t = 0_J \qquad (95)$$

$$\nabla_m \phi^t(p^t,v^t,x^t,m^t) - e^t \omega^t = 0_K$$

$$\rho^t \cdot x^t - w^t \cdot m^t + b^t = 0, \qquad t = 0,1$$

Assumptions (ii), (iii) and (v) imply that x^t,m^t satisfies Samuelson's [1947] strong second order sufficiency conditions for the tth problem in (91), and hence once differentiable solution functions $x^t(\rho,\omega,b)$ and $m^t(\rho,\omega,b)$ for the problems (91) exist [35] for ρ,ω,b near

(ρ^t,ω^t,b^t) for $t = 0,1$. Note that $g^t(p^t,v^t\rho,\omega,b) = \phi^t(p^t,v^t,x^t(\rho,\omega,b), m^t(\rho,\omega,b)) = \phi^t(p^t,v^t,x^t(\rho,\omega,b), m^t(\rho,\omega,b)) + e^t(\rho,\omega,b)[\rho\cdot x^t(\rho,\omega,b) - \omega\cdot m^t(\rho,\omega,b) + b]$ for (p,ω,b) close to (ρ^t,ω^t,b^t). Differentiate this last equation with respect to the components of ρ,ω,b, evaluate the resulting derivatives at ρ^t,ω^t,b^t, use the first order conditions (95), and we have the relations (92)-(94). The positivity of e^t follows from (92), $x^t >> 0_J$ and assumption (v). Q.E.D.

Corollary: If in addition to the hypotheses of the theorem, $\phi^t(p,v,x,m)$ is twice continuously differentiable with respect to all of its arguments in a neighbourhood of (p^t,v^t,x^t,m^t) for $t = 0,1$, then the first order derivatives of g^t with respect to p and v satisfy the following relations:

$$\nabla_p g^t(p^t,v^t,\rho^t,\omega^t,b^t) = \nabla_p \phi^t(p^t,v^t,x^t,m^t) = y^t; \tag{96}$$

$$\nabla_v g^t(p^t,v^t,\rho^t,\omega^t,b^t) = \nabla_v \phi^t(p^t,v^t,x^t,m^t) = w^t;\ t = 0,1 \tag{97}$$

where w^t is the vector of period t domestic input prices for the economy.

Proof: The first set of equalities in (96) and (97) are derived using the Implicit Function Theorem in the same manner as we derived (93)-(94) in proving the theorem. The second set of equalities in (96) and (97) are derived using known duality theory results; e.g., see Diewert, [1974; pp.137-140]. Q.E.D.

Note that (94) yields an economic interpretation for the period t Lagrange multiplier e^t: it is the marginal increase in domestic output that will occur if we allow the balance of trade deficit to increase by a marginal unit; i.e., it is the economy's **exchange rate** for period t, giving the number of domestic dollars that can be produced by a gift of one foreign dollar.[36]

Note also that the right-hand sides of (92)-(94) and (96)-(97) are all observable in principle. Furthermore, note that the economy's system of export supply functions $x^t(p,v,\rho,\omega,b)$

may be obtained by taking the ratio of (92) to (94), the economy's system of import demand functions $m^t(p,v,\rho,\omega,b)$ may be obtained by taking minus the ratio of (93) to (94),[37] the economy's system of domestic supply functions $y^t(p,v,\rho,\omega,b)$ is given by (96) and the economy's system of competitive domestic input prices $w^t(p,v,\rho,\omega,b)$ is given by (97). All of these functions (and the exchange rate) can be obtained by differentiating the period t balance of trade restricted value added function $g^t(p,v,\rho,\omega,b)$.

We are finally ready to define some approximations to the terms of trade adjustment indexes $\bar{A}^0$ and $\bar{A}^1$ defined in (87). Using (86) and (87) and assuming the hypotheses of Theorem 14, we have

$$
\begin{aligned}
\bar{A}^0 &\equiv g^0(p^0,v^0,\rho^1\omega^1,b^1)/g^0(p^0,v^0,\rho^0,\omega^0,b^0) \\
&= g^0(p^0,v^0,\bar{\rho}^0,\omega^1,b^1)/p^0\cdot y^0 \\
&\cong [g^0 + \nabla_\rho g^0\cdot(\rho^1-\rho^0) + \nabla_\omega g^0\cdot(\omega^1-\omega^0) + \nabla_b g^0\cdot(b^1-b^0)]/p^0\cdot y^0 \\
&\qquad \text{making a first order Taylor series approximation} \\
&= [p^0\cdot y^0 + e^0x^0\cdot(\rho^1-\rho^0) - e^0m^0\cdot(\omega^1-\omega^0) + e^0(b^1-b^0)]/p^0\cdot y^0 \\
&= 1 + e^0[x^0\cdot(\rho^1-\rho^0) - m^0\cdot(\omega^1-\omega^0) + b^1-b^0]/p^0\cdot y^0 \\
&\equiv a^0(\rho^0,\rho^1,x^0,x^1,\omega^0,\omega^1,m^0,m^1,e^0,e^1,b^0,b^1,p^0,p^1,y^0,y^1). \qquad (98)
\end{aligned}
$$

Thus the function a^0 defined by (98) is a first order approximation to the theoretical terms of trade adjustment factor $\bar{A}^0$. Note that the function a^0 may be evaluated if we can observe prices and quantities for the private sector of the economy for periods 0 and 1.

In similar fashion, we find that a^1 defined by (99) below is a first order approximation

to the theoretical terms of trade adjustment factor $\bar{A}^1$:

$$\bar{A}^1 \equiv g^1(p^1,v^1,\rho^1,\omega^1,b^1)/g^1(p^1,v^1,\rho^0,\omega^0,b^0)$$

$$= p^1\cdot y^1/g^1(p^1,v^1,\rho^0,\omega^0,b^0)$$

$$\cong \left\{1 - e^1[x^1\cdot(\rho^1-\rho^0) - m^1\cdot(\omega^1-\omega^0) + b^1-b^0]/p^1\cdot y^1\right\}^{-1}$$

$$\equiv a^1(\rho^0,\rho^1,x^0,x^1,\omega^0,\omega^1,m^0,m^1,e^0,e^1,b^0,b^1,p^0,p^1,y^0,y^1). \quad (99)$$

In the previous three sections, we were always able to obtain Paasche and Laspeyres type bounds for our theoretical indexes. In the present context, we do not seem to be able to find these bounds. However, in place of bounds, we are able to obtain the (first order) approximating functions[38] a^0 and a^1 for the theoretical indexes $\bar{A}^0$ and $\bar{A}^1$. Furthermore, we shall approximate the geometric mean of $\bar{A}^0$ and $\bar{A}^1$ by the geometric mean of a^0 and a^1; i.e.,

$$(\bar{A}^0\bar{A}^1)^{\frac{1}{2}} \cong [a^0(\rho^0,\ldots,y^1)a^1(\rho^0,\ldots,y^1)]^{\frac{1}{2}}$$

$$\equiv a_G(\rho^0,\rho^1,x^0,x^1,\omega^0,\omega^1,m^0,m^1,e^0,e^1,b^0,b^1,p^0,p^1,y^0,y^1). \quad (100)$$

By how much will the index number formulae for a^0, a^1 and a_G differ? The following theorem indicates that for "normal" economic data that come from adjacent time periods, so that differences between prices and quantities in the two periods are not too large, the three numerically implementable terms of trade adjustment indexes will all give about the same answer.

Theorem 15: Let the functions a^0, a^1 and a_G be defined by (98), (99) and (100). Then these three functions differentially approximate each other to the first order around any point where the prices and quantities of the two periods are equal; i.e., we have $a^0 = a^1 = a_G$ and $\nabla a^0 = \nabla a^1 = \nabla a_G$ where ∇ stands for the vector of partial derivatives of the function with respect to all components of the vector $(\rho^0,\rho^1,x^0,x^1,\omega^0,\omega^1,m^0,m^1,e^0,e^1,b^0,b^1,p^0,p^1,y^0,y^1)$ and each function is evaluated at a common point where $\rho^0 = \rho^1$, $x^0 = x^1$, $\omega^0 = \omega^1$, $m^0 = m^1$, $e^0 = e^1$, $b^0 = b^1$, $p^0 = p^1$, and $y^0 = y^1$.

Theorem 15 may be proven by a series of straightforward but lengthy computations.

As usual, we can develop a translog approach to the problem of finding empirically implementable approximations to the theoretical terms of trade adjustment factor $(\bar{A}^0,\bar{A}^1)^{1/2}$. Assume that the economy's period t balance of trade restricted value added function g^t has the following translog functional form: for $(p,v) \equiv q >> 0_{M+N}$ and $(\rho,\omega,b) \equiv z >> 0_{J+K+1}$,

$$\ell n g^t(q,z) \equiv \alpha_0^t + \sum_{i=1}^{M+N} \alpha_i^t \ell n q_i + \frac{1}{2}\sum_{i=1}^{M+N}\sum_{j=1}^{M+N} \alpha_{ij}^t \ell n q_i \ell n q_j + \sum_{i=1}^{M+N}\sum_{j=1}^{J+K+1} \gamma_{ij}^t \ell n q_i \ell n z_j \tag{101}$$

$$+ \sum_{i=1}^{J+K+1} \beta_i^t \ell n z_i + \frac{1}{2}\sum_{i=1}^{J+K+1}\sum_{j=1}^{J+K+1} \beta_{ij}^t \ell n z_i \ell n z_j; \qquad t = 0,1,$$

where $\alpha_{ij}^t = \alpha_{ji}^t$ and $\beta_{ij}^t = \beta_{ji}^t$ for all i, j and t.

Theorem 16: Assume the hypotheses of Theorem 14.[39] In addition, assume: (i) g^t has the translog functional form defined by (101) for $t = 0,1$, (ii) $b^0 > 0$ and $b^1 > 0$, and (iii) the translog functional forms for g^0 and g^1 have identical quadratic coefficients for the terms involving $z \equiv (\rho,\omega,b)$; i.e., $\beta_{ij}^0 = \beta_{ij}^1$ for $1 \leq i, j \leq J+K+1$. Define the theoretical terms of trade adjustment factors $\bar{A}^0$ and $\bar{A}^1$ by (87). Then we have the following empirically

computable exact expression for the geometric mean of these theoretical indexes:

$$[\bar{A}^0\bar{A}^1]^{\frac{1}{2}} = \frac{\Pi_{j=1}^{J}(\rho_j^1/\rho_j^0)^{\frac{1}{2}[\frac{e^0\rho_j^0x_j^0}{p^0\cdot y^0}+\frac{e^1\rho_j^1x_j^1}{p^0\cdot y^0}]}(b^1/b^0)^{\frac{1}{2}[\frac{e^0b^0}{p^0\cdot y^0}+\frac{e^1b^1}{p^1\cdot y^1}]}}{\Pi_{k=1}^{K}(\omega_k^1/\omega_k^0)^{\frac{1}{2}[\frac{e^0\omega_k^0m_k^0}{p^0\cdot y^0}+\frac{e^1\omega_k^1m_k^1}{p^0\cdot y^0}]}} \tag{102}$$

$$\equiv a_T(\rho^0,\rho^1,x^0,x^1,\omega^0,\omega^1,m^0,m^1,e^0,e^1,b^0,b^1,p^0,p^1,y^0,y^1).$$

The proof of the theorem is a straightforward application of the translog identity in Caves, Christensen and Diewert [1982; p.1412]. Note that the right-hand side of (102), which we define to be the translog terms of trade adjustment factor a_T, is essentially an index number of export prices times a balance of trade deficit index number divided by an index number of import prices. Hence if export prices rise relative to import prices, a_T will be greater than unity. Moreover, a_T will give us the approximate proportion of private production sector growth that can be attributed to the improvement in the terms of trade. Note also that if export and import prices are constant but the deficit increases so that $b^1 > b^0 > 0$, then again $a_T > 1$.

The derivation of (102) required the positivity of the period 0 and 1 balance of trade deficits; i.e., we required $b^0 > 0$ and $b^1 > 0$. However, if $b^1 = b^0 = 0$, we may still apply the translog identity in Caves, Christensen and Diewert and again we will obtain (102), except the last term involving (b^1/b^0) is dropped from the right-hand side of (102). If b^1 and b^0 are both negative, then it turns out that we may still derive formula (102). However, if b^1 and b^0 are of opposite sign, then it appears that we cannot derive (102).[40]

We have now found four empirically implementable terms of trade adjustment indexes, a^0, a^1, a_G and a_T defined by (98), (99), (100) and (102) respectively. How close will they

be in practice? The following theorem says that they will all be reasonably close but a_G and a_T will be very close.

Theorem 17: The functions a^0, a^1, a_G and a_T differentially approximate each other to the first order around any point where the prices and quantities pertaining to period 0 equal the corresponding prices and quantities pertaining to period 1. Moreover, a_G and a_T differentially approximate each other to the second order around any point where the prices and quantities pertaining to period 0 equal the corresponding prices and quantities pertaining to period 1; i.e., $a_G = a_T$, $\nabla a_G = \nabla a_T$ and $\nabla^2 a_G = \nabla^2 a_T$ where ∇^2 stands for the matrix of second order partial derivatives of the function with respect to all components of $(\rho^0,\rho^1,x^0,x^1,\omega^0,\omega^1,m^0,m^1,e^0,e^1,b^0,b^1,p^0,p^1,y^0,y^1)$ and each function is evaluated at a common point where $\rho^0 = \rho^1$, $x^0 = x^1$, $\omega^0 = \omega^1$, $m^0 = m^1$, $e^0 = e^1$, $b^0 = b^1$, $p^0 = p^1$ and $y^0 = y^1$. Each price and quantity is assumed to be positive, except b^0 and b^1 have the same sign (which could be negative).

The proof of Theorem 17 requires a horrendous number of straightforward computations.

In Sections 2-4, we developed both nonparametric bounding and translog approaches in order to obtain good empirical approximations to various theoretical indexes. In each case, we found that the two approaches led to empirical indexes that approximated each other to the second order. In the present section, the nonparametric bounding approach has been replaced by a nonparametric first order approximation approach. Again, we find that the nonparametric and the translog approach led to empirical terms of trade adjustment indexes (a_G and a_T) that approximate each other to the second order. We recommend that either of the indexes a_G or a_T be used in empirical applications.[41]

We conclude this section by returning to the topic that we treated in the previous section, namely productivity indexes. In Section 4, we derived two good measures of the increase in efficiency of the economy, H_T defined by (65) and H_F defined by (66). In the light of the analysis presented in the present section, it can seen that H_T and H_F incorporate not only the efficiency growth in the economy between periods 0 and 1, but also output growth due to changes in the terms of trade. Hence in order to obtain empirical measures of pure efficiency growth, H_T and H_F should be divided by a_T or a_G; i.e., define

the **pure efficiency indexes** E_T and E_F by

$$E_T \equiv H_T/a_T;\ E_F \equiv H_F/a_G. \qquad (103)$$

E_T and E_F may also be regarded as terms of trade adjusted indexes of total factor productivity.[42]

6. Practical Problems with the Theory

The empirical approximations we have suggested for the various theoretical indexes defined in the previous four sections depend on ''observable'' prices and quantities. Just how observable are these prices and quantities? Moreover, we have made a large number of auxiliary assumptions in the proofs of our approximation theorems. Just how restrictive are these assumptions in practice?

We see at least four main classes of difficulties with our approximation theorems, which may be grouped as follows: (i) the existence of too many goods, (ii) the existence of durable inputs (capital), (iii) the existence of taxes, and (iv) the existence of noncompetitive behaviour on the part of producers. We discuss each source of difficulty in turn.

The existence of too many goods. We have developed various formulae that require price and quantity information on M outputs, N inputs, J exports and K imports. This is fine as far as it goes, but we must recognize that M, N, J and K can easily be in the millions in a modern Western economy. We are **never** going to be able to collect accurate information on all of these goods. Hence price and quantity information will have to be collected on a selective basis and various simplifying assumptions about the unobserved prices and quantities will have to made. This will reduce the accuracy of our approximation formulae in an essentially unknown manner.

The existence of durable inputs. If there are inputs which last longer than the period (such as machines, structures, goods in process, inventories), then the leftover portion should be added to the economy's list of outputs. These end of the period leftover components of the capital stock will become components of next periods' starting capital stock. This

view of capital is associated with the names of von Neumann [1945-46], Hicks [1946; pp.323-328] and Malinvaud [1953]. The problem with all of this is that usually, there are no active second-hand markets for these leftover capital goods. Hence, we are forced to make guesses about these prices. An alternative to treating capital as both an input and an output is to treat it as an input that deteriorates a certain amount (independent of use)[43] over the period and then construct the expected **user cost** for the capital good. Consider a durable good that can be purchased at the beginning of period 0 at the observed period 0 price q^0. The producer could use the capital good during period 0 and then sell his used durable goods next period at a (possibly hypothetical) second-hand market at an **expected** price $\tilde{q}^{01}$. Assuming that the producer can borrow or lend at the interest rate r, we may follow Hicks [1946] and compute the present value of the cost of buying one unit of the good, using it for one period and selling it next period. The resulting period 0 user cost w^0 is

$$w^0 \equiv q^0 - \tilde{q}^{01}/(1+r) = [q^0 r + (q^0 - \tilde{q}^{01})]/(1+r). \qquad (104)$$

The first term on the right-hand side of (104) is an interest cost term while the second term combines the effects of depreciation and anticipated capital gains. We can separate out these two effects if we let $\tilde{q}^1$ be the producer's **expected** spot price for a new unit of the durable during period 1. Then we may rewrite (104) as

$$w^0 = [q^0 r + (\tilde{q}^1 - \tilde{q}^{01}) - (\tilde{q}^1 - q^0)]/(1+r) \qquad (105)$$

The above derivation of the user cost formulae (104) and (105) parallels a similar derivation made in the consumer context by Diewert [1983].

We could similarly derive a formula for the user cost of the durable for period 1, w^1 say, and then w^0 and w^1 would be appropriate input prices for the formulae that appear in Section 4 above.

It all sounds very straightforward, but there are severe difficulties: (i) there may be no second-hand market for the durable and (ii) even if there is a second-hand market for the durable, it is **not** the actual second-hand market price that appears in (104), but the producer's **expectation** of it, $\tilde{q}^{01}$. In periods of rapid inflation, producer's expectations of

future asset prices may be wildly inaccurate. Some further difficulties with (104) are: (iii) it may be difficult to determine what the relevant interest rate r is and (iv) the different income tax treatment of different classes of capital goods creates further complications.

Hicks [1981; Ch. 8, 9 and 11] provides perhaps the best overall discussion of the severe measurement problems that the existence of capital goods (along with the nonexistence of second-hand markets) creates. Unfortunately, he does not present any firm recommendations on the solution of these problems.[44]

These measurement problems are not unimportant. For example, whenever oil prices increase dramatically, the asset prices of certain fuel inefficient cars, trucks and planes drops dramatically. Hence the user cost of these cars, trucks and planes will dramatically increase during this period. However, conventional national income accounting procedures tend to ignore capital gains and losses, and so conventionally measured total factor productivity will tend to show a ''puzzling'' decline, which may well disappear if we could construct the correct user costs.

The existence of taxes. From the viewpoint of production theory, the treatment of indirect taxes is straightforward: any sales taxes, excise taxes, property taxes, tariffs, duties, unemployment insurance premiums, pension payments, etc., that are actually a cost to the producer should be included in the corresponding input price, but taxes which fall on the outputs of producers should **not** be included in the corresponding output prices. However, subsidies that producers receive on the sales of their outputs should be included in the corresponding output prices. If a good that is produced by one industry is taxed and then used as an input into other industries, then this good should be considered as both an output (which sells at the before tax price) and as an input (which must be purchased at the after tax price).[45] Unfortunately, the conventional national income accounting treatment of indirect taxes is not usually consistent with the above suggested treatment: all indirect taxes tend to be lumped together no matter whether they fall on inputs or outputs. The existence of nonneutral corporate or noncorporate income taxes also creates practical measurement difficulties. The problem is that in most Western economies, the business income tax allows a conventional accounting deduction for the depreciation of an asset which is not equal to the true economic depreciation. This makes the user cost

formulae (104) or (105) more complicated. Moreover, many Western countries allow capital assets in different industries to be treated differently from a tax viewpoint, which leads to deadweight loss[46] and even more complicated user cost formulae.

The existence of noncompetitive behaviour. Much of our theoretical analysis and virtually all of our "practical" approximations rest on the assumption that each producer competitively maximizes profits.[47] However, there are whole branches of economics that assume that producers do not behave competitively. At the present time, the only type of noncompetitive behaviour that our methodology can treat is markup monopolistic behaviour, provided that we know the markup. In this case, the markup can be treated as a tax that must be deducted from the observed selling price of the output. However, in practice, how are we going to determine the markup? It must be conceded that the widespread existence of noncompetitive behaviour limits the usefulness of our methodology.

7. Conclusion

Our recommended measures of output price change (inflation) from the production theory perspective are P_T and P_F defined in Section 2, our recommended measures of real output growth are Q_T (see 54) and Q_F defined in Section 3, and our recommended measures of efficiency growth are H_T and H_F (see (65) and (66)) and E_T and E_F (see (103)). The efficiency growth measures H_T and H_F include any contribution due to changes in the terms of trade between the two periods under consideration, while E_T and E_F are "pure" efficiency measures that exclude any contribution due to changes in the terms of trade.

We noted in the introduction that aggregate consumer based measures of welfare change do not command popular acceptance due to the difficulties involved in weighting the relative welfares of different consumer groups. However, in Section 6, we noted that producer oriented measures of welfare change (such as H_T or E_T) are also subject to some difficulties. Still, the producer oriented measures of welfare change seem more promising to us at this point.

It may be that the most important contribution of this paper occurs in Section 5, where we derive the terms of trade adjustment indexes a_G and a_T (see (100) and (102)) which allow us to estimate what proportion of an economy's private production sector growth between two periods is due to changes in the international trade environment facing the economy.

Footnotes

* This research was supported by Statistics Canada and the SSHRC of Canada. Neither institution is responsible for the views expressed here.

1 This consumer oriented approach was developed by Pigou [1920], Hicks [1939, 1940, 1975] and Samuelson [1950] on the quantity side and by Konüs [1924] and Pollak [1971, 1981] on the price side.

2 This producer oriented approach was developed by Hicks [1940, 1975, 1981], Samuelson [1950, 1961], Bergson [1961] and Moorsteen [1961] on the quantity side and by Fisher and Shell [1972], Samuelson and Swamy [1974] and Archibald [1977] on the price side.

3 The first (vi) points are listed in Hicks [1975; p.317] and point (vii) may be found in Hicks [1939] or [1940].

4 Also the first part of the theory of productivity indexes is applicable to planned economies as well as to market economies, a point noted by Hicks [1978; p.318].

5 See Samuelson [1953-54; p.20]. Alternative terms for this concept include (i) the gross profit function (Gorman [1968]), (ii) the restricted profit function (McFadden [1978] and Lau [1976]), (iii) the variable profit function (Diewert [1973, 1974]), (iv) the (nominal) value added function (Khang [1971], Sato [1976; p.438] and Bruno [1978]) and (v) the revenue function (McFadden [1978] and Archibald [1977]). The above references derive the properties of π^t under various alternative assumptions about S^t. We shall require that the maximum in (i) exist. This will be the case if for every $v \geq 0_N$, $\left\{ y: (y,v) \epsilon S^t \right\}$ is closed and bounded. Notation: $v >> 0_N$ means each component is nonnegative, and $v > 0_N$ means $v \neq 0_N$ and $v \geq 0_N$. 1_N is a vector of ones.

6 A set of sufficient conditions is that S^0 and S^1 both satisfy the following conditions: (i) S^t is a nonempty, closed subset of the nonnegative orthant in R^{M+N} (hence the role of inputs and outputs cannot change), (ii) if $v >> 0_N$, then there exists $y >> 0_M$ such that $(y,v)\epsilon S^t$ (positive amounts of all inputs yield positive amounts of all outputs), (iii) $v \geq 0_N$, $(y,v)\epsilon S^t$ implies that there exists a positive scalar k(v) that can depend on the input vector v such that $y \leq k(v)1_M$ (bounded inputs produce bounded outputs), (iv) if $(y^1,v^1)\epsilon S^t$, $v^2 \geq v^1$, $0_M \leq y^2 \leq y^1$, then $(y^2,v^2)\epsilon S^t$ (free disposal), and (v) $(y^1,v^1)\epsilon S^t$, $(y^2,v^2)\epsilon S^t$, $0 \leq \lambda \leq 1$, then $(\lambda y^1 + (1-\lambda)y^2, \lambda v^1 + (1-\lambda)v^2)\epsilon S^t$ (convexity of S^t). In what follows, we call the above regularity conditions on the technology set S^t **Conditions I**. The proof of continuity in Theorem 3 requires the use of the Debreu [1952; pp.889-890, 1959; p.19] and Berge [1963; p.116] Maximum Theorem.

7 Hence we recommend the use of the chain principle rather than using a fixed base when constructing an output price index.

8 Our definition is perhaps slightly different than the Fisher-Shell definition, but it captures the same idea.

9 Our definition is somewhat more general than that of Samuelson and Swamy who assume only one input and no technical change. Sato [1976; p.438] has the many input definition without technical change. Both sets of authors noted the analogy of the output index defined by (23) to the Allen [1949] utility or real income index in the context of consumer theory.

10 If the minimum does not exist, replace the minimum by an infinum.

11 Malmquist developed the basic idea in the context of consumer theory and the idea was explicitly used by Bergson [1961] and Moorsteen [1961] in the context of producer theory. Other authors who have used the basic concept involved in this index include Samuelson and Swamy [1974; pp.590-591], Hicks [1981; p.256] and Caves, Christensen and Diewert [1982; pp.1399-1401].

12 See also Diewert [1981; p.178].

13 A function f is homothetic if and only if it is a monotonically increasing transform of a positively linearly homogeneous function g; i.e., $f(x) = F(g(x))$ for some monotonically increasing function of one variable F.

14 For references to the literature on separability concepts, see Blackorby, Primont and Russell [1978].

15 See Gorman [1968] or Diewert [1973; p.293].

16 In order to establish this equality, we need to use the positive linear homogeneity of $\pi^t(p,v)$ in p for fixed v. If S^t satisfies Conditions I listed in Footnote 6, then π^t satisfies the following Conditions II (and vice versa): (i) $\pi^t(p,v)$ is a nonnegative continuous function defined for all $p \geq 0_M$ and $v \geq 0_N$, and is positive if $p > 0_M$ and $v >> 0_N$, (ii) $\pi^t(p,v)$ is positively linearly homogeneous, nondecreasing and convex in p for each fixed $v \geq 0_N$, and (iii) $\pi^t(p,v)$ is concave and nondecreasing in v for each fixed $p \geq 0_M$.

17 Woodland [1978] shows how to set up an econometric test for this kind of separability.

18 The assumptions $y^0 > 0_M$ and $y^1 > 0_M$ can be relaxed to $p^0 \cdot y^0 > 0$ and $p^1 \cdot y^1 > 0$. This is relevant if there are intermediate inputs in the list of outputs.

19 Of course, if S^0 and S^1 satisfy the homothetic separability assumption (29) or (39), then $Q_L \leq Q^0(y^0,y^1,v) = Q^1(y^0,y^1,v) \leq Q_P$ for any reference input vector $v >> 0_N$.

20 If there is not technical progress so that $S^0 = S^1$, then of course $d(y,v,t) = d^0(y,v) = d^1(y,v)$ for all t.

21 A set of sufficient conditions is that S^0 and S^1 satisfy Conditions I listed in Footnote 6 and $v^0 >> 0_N$, $v^1 >> 0_N$. It should be noted that if S^t satisfies Conditions I then d^t defined by (25) will satisfy the following Conditions III: (i) $d^t(y,v)$ is a positive continuous function defined for all $y >> 0_M$ and $v >> 0_N$, (ii) for $\lambda > 0$, $d^t(\lambda y,v), = \lambda d^t(y,v)$, (iii) if $0_M << y^1 \leq y^2$, $0_N << v^2 \leq v^1$, then $d^t(y^1,v^1) \leq d^t(y^2,v^2)$, and (iv) if $0_M << y^i$, $0_N << v^i$ for $i = 1,2$, $0 < \lambda < 1$, then $d^t(\lambda d^t(y^1,v^1)^{-1}y^1 + (1-\lambda)d^t(y^2,v^2)^{-1}y^2, \lambda v^1 + (1-\lambda)v^2)^{-1} \geq 1$. Properties (i), (ii) and (iv) imply that $(d^t)^{-1}$ is a quasiconcave function in (y,v) and is concave in v for each fixed y. Since $d^t(y,v)^{-1}$ is positive, quasiconcave and linearly homogeneous in y for each fixed v, by a result due to Berge [1963; 208], $d^t(y,v)^{-1}$ is a concave function of y for each fixed v.

22 This concept was called the Malmquist output based productivity index in Caves, Christensen and Diewert [1982; p.1402], but the concept appears to be due to Hicks [1961, 1981; p.256] who called it an efficiency index.

[23] A sufficient condition for this is: S^0 and S^1 satisfy Conditions I listed in Footnote 6.

[24] Theorem 12 corresponds to case (ii) and Corollary 12.1 corresponds to case (i) in Caves, Christensen and Diewert [1982; p.1406]. The returns to scale term defined by (47) in Caves, Christensen and Diewert has a typographical error and should be replaced by the reciprocal of the right-hand side of (47).

[25] We have used the subscript F for Irving Fisher [1922] since in the zero profits case, $H_F = Q_F(p^0,p^1,y^0,y^1)/Q_F(w^0,w^1,v^0,v^1)$, the geometric mean of the Paasche and Laspeyres output indexes divided by the geometric mean of the Paasche and Laspeyres input indexes.

[26] These input growth terms could be obtained by dividing any of the quantity indexes developed in Section 3 by the productivity indexes developed in Section 4.

[27] Let (y,x) replace y and let (v,m) replace v. Then Conditions I become Conditions IV.

[28] If there is no y such that $(y,v,x,m)\epsilon T^t$ then define $\phi^t(p,v,x,m) = -\infty$. This case could arise if x is chosen to be too big relative to v and m, so that the economy is unable to produce the export vector x given the available domestic input vector v and import vector m.

[29] The main change going from Conditions II to Conditions V is that $\phi^t(p,v,x,m)$ is nondecreasing in the components of v and m, and nonincreasing in the components of x.

[30] The export prices include any producer subsidies and the import prices include any tariffs and domestic border taxes that may be levied on imports; i.e., the "foreign" price vectors ρ and ω are the after tax, tariff and subsidy prices that domestic producers face, denominated in units of a foreign numeraire currency.

[31] If $b < 0$, then $-b$ is the size of the surplus that the private producer sector must earn on international transactions for the period.

[32] If there is no y, x, m such that $(y,v,x,m)\epsilon T^t$ and $\rho\cdot x - \omega\cdot m + b \geq 0$, then $g^t(p,v,p,\omega,b) \equiv -\infty$.

[33] $g^t(p,v,\rho,\omega,b)$ will be: (i) nondecreasing and concave in v, b for fixed p, ρ, w, (ii) positively linearly homogeneous, nondecreasing and convex in p for fixed v, ρ, ω, b, (iii) positively homogeneous of degree zero and quasiconvex in p, ω, b for fixed p, v, (iv) nondecreasing in v, p for fixed ρ, ω, b and (v) nonincreasing in the components of ω for fixed p, v, ρ and b. A complete dual characterization of g^t is technically quite involved and will not be attempted here.

[34] These inequalities are essentially due to Archibald [1977; p.62].

[35] See also Diewert and Woodland [1977; pp.391-396]. Differentiate equations (95) with respect to x, m, e, ρ, ω and b and use the Implicit Function Theorem to solve for x, m, e as functions of ρ, ω and b. Assumption (iii) implies the appropriate bordered Hessian matrix is nonsingular.

[36] This economic interpretation of Lagrange multipliers is due to Samuelson [1947; p.65 and p.132].

[37] These relationships are the production theory counterparts to Woodland's [1980; p.919] Identities, derived in a general equilibrium context.

38 Note that a^0 may be rewritten as a function involving the Laspeyres export and import prices indexes as follows: $a^0 = 1 + e^0[(\rho^1\cdot x^0/p^0\cdot x^0)-1)]\ (\rho^0\cdot x^0/p^0\cdot y^0) - e^0[(\omega^1\cdot m^0/\ \omega^0\cdot m^0)-1]\cdot (\omega^0\cdot m^0/\ {}^0\cdot y^0) + e^0\,(b^1 - b^0)/p^0\cdot y^0$. Similarly, a^1 may be rewritten as a function involving the Paasche export and import price indexes $\rho^1\cdot x^1/\rho^0\cdot x^1$ and $\omega^1\cdot m^1/\omega^0\cdot m^1$.

39 We need to be able to use the derivative relations (92)-(94). These relations can be established under weaker hypotheses using duality theory techniques, a generalization that we do not attempt here due to space limitations.

40 Note that b^0 and b^1 do not have to be sign restricted in the definitions of a^0, a^1 and a_G.

41 In place of $a_G \equiv (a^0a^1)^{1/2}$, we could use $\frac{1}{2}a^0 + \frac{1}{2}a^1$ or any other symmetric mean of a^0 and a^1. These symmetric means of a^0 and a^1 differentially approximate a_G and a_T to the second order as in Theorem 17; see Diewert [1978; p.897] for the definition of a symmetric mean and an indication of how to prove this approximation assertion.

42 The indexes E_T and E_F are not the only empirically implementable trade adjusted indexes of efficiency that could be derived; however, we conjecture that all "good" indexes will differentially approximate E_T or E_F to the second order around an equal price and quantity point.

43 Hicks [1946; p.326] indicates how to model the general case where deterioration can depend on how intensively the capital good is used.

44 The only partial solution that I can offer is that we use rental market information whenever possible; i.e., many assets are leased and so their rental price w^0 may be calculated directly using market data. It may be necessary to extrapolate these rental prices to other assets which are not leased.

45 It should be noted that such taxes that are internal to the production sector of the economy are extremely inefficient and should not be used by a rational government.

46 See Diewert [1981b] on the measurement of the deadweight loss due to nonneutral corporate income taxes and for references to the literature.

47 This assumption also allows us to aggregate over sectors without any difficulty; see Gorman [1968] and Diewert [1980; pp.464-467].

References

Archibald, R.B. [1977], "On the Theory of Industrial Price Measurement: Output Price Indexes", *Annals of Economic and Social Measurement* 6, pp.57-72.

Allen, R.G.D. [1949], "The Economic Theory of Index Numbers", *Economica* 16, pp.197-203.

Berge, C. [1963], *Topological Spaces*, New York: Macmillan.

Bergson, A. [1961], *National Income of the Soviet Union Since 1928*, Cambridge, Mass.: Harvard University Press.

Blackorby, C., D. Primont and R.R. Russell [1978], *Duality, Separability and Functional Structure: Theory and Economic Applications*, New York: American Elsevier.

Bruno, M. [1978], "Duality, Intermediate Inputs and Value Added", in *Production Economics: A Dual Approach to Theory and Application*, Vol. 2, M. Fuss and D. McFadden (eds.), Amsterdam: North-Holland.

Caves, D.W., L.R. Christensen and W.E. Diewert [1982], "The Economic Theory of Index Numbers and the Measurement of Input, Output, and Productivity", *Econometrica* 50, pp.1393-1414.

Debreu, G. [1952], "A Social Equilibrium Existence Theorem", *Proceedings of the National Academy of Sciences* 38, pp.886-893.

_________ [1959], *Theory of Value*, New York: John Wiley.

Diewert, W.E. [1973], "Functional Forms for Profit and Transformation Functions", *Journal of Economic Theory* 6, pp.284-316.

__________ [1974], ''Applications of Duality Theory'', pp. 106-171 in *Frontiers of Quantitative Economics*, Vol. II, M.D. Intriligator and D.A. Kendrick (eds.), Amsterdam: North-Holland.

__________ [1978], ''Superlative Index Numbers and Consistency in Aggregation'', *Econometrica* 46, pp.883-900.

__________ [1980], ''Aggregation Problems in the Measurement of Capital'', pp. 433-528 in *The Measurement of Capital*, D. Usher (ed.), Chicago: The University of Chicago Press.

__________ [1981], ''The Economic Theory of Index Numbers: A Survey'', pp. 163-208 in *Essays in the Theory and Measurement of Consumer Behaviour in Honour of Sir Richard Stone*, A. Deaton (ed.), London: Cambridge University Press.

__________ [1981b], ''On Measuring the Loss of Output Due to Nonneutral Business Taxation'', pp.57-80 in *Depreciation, Inflation and the Taxation of Income from Capital*, C.R. Hulten (ed.), Washington D.C.: Urban Institute Press.

__________ [1982], ''Duality Approaches to Microeconomic Theory'', pp.535-599 in *Handbook of Mathematical Economics* Vol. II, K.J. Arrow and M.D. Intriligator (eds.), Amsterdam: North-Holland.

__________ [1983], ''The Theory of the Cost-of-Living Index and the Measurement of Welfare Change'', *Price Level Measurement: Proceedings From a Conference Sponsored by Statistics Canada.*

__________ and A.D. Woodland [1977], ''Frank Knight's Theorem in Linear Programming Revisited'', *Econometrica* 45, pp.375-398.

Fisher, I. [1922], *The Making of Index Numbers*, Boston: Houghton Mifflin Co.

Fisher, F.M. and K. Shell [1972], ''The Pure Theory of the National Output Deflator'', pp.49-113 in *The Economic Theory of Price Indexes* by F.M. Fisher and K. Shell, New York: Academic Press.

Gorman, W.M. [1968], ''Measuring the Quantities of Fixed Factors'', pp.141-172 in *Value, Capital and Growth: Papers in Honour of Sir John Hicks*, J.N. Wolfe (ed.), Chicago: Aldine.

Hicks, J.R. [1939], ''The Foundation of Welfare Economics'', *Economic Journal* 49, pp.696-172.

——— [1940], ''The Valuation of the Social Income'', *Economica* 7, pp.105-124.

——— [1961], ''Measurement of Capital in Relation to the Measurement of Other Economic Aggregates'', in F.A. Lutz and D.C. Hague (eds.), *The Theory of Capital*, London, Macmillan.

——— [1975], ''The Scope and Status of Welfare Economics'', *Oxford Economic Papers* 27, pp.307-326.

Hicks, J. [1981], *Wealth and Welfare*, Cambridge, Mass.: Harvard University Press.

Khang, C. [1971], ''An Isovalue Locus Involving Intermediate Goods and its Applications to the Pure Theory of International Trade'', *Journal of International Economics* 1, pp.315-325.

Konüs, A.A. [1924], ''The Problem of the True Index of the Cost-of-Living'', translated in *Econometrica* 7, [1939], pp.10-29.

Kohli, U.R.J. [1978], ''A Gross National Product Function and the Derived Demand for Imports and Supply of Exports'', *Canadian Journal of Economics* 11, pp.167-182.

Lau, L.J. [1976], ''A Characterization of the Normalized Restricted Profit Function'', *Journal of Economic Theory* 12, pp.131-163.

Malinvaud, E. [1953], "Capital Accumulation and the Efficient Allocation of Resources", *Econometrica* 21, pp.233-268.

Malmquist, S. [1953], "Index Numbers and Indifference Surfaces", *Trabajos de Estadistica* 4, pp.209-242.

McFadden, D. [1978], "Cost, Revenue and Profit Functions", pp.3-109 in *Production Economics: A Dual Approach to Theory and Applications*, Vol. I, M. Fuss and D. McFadden (eds.), Amsterdam: North-Holland.

Moorsteen, R.H. [1961], "On Measuring Productive Potential and Relative Efficiency", *Quarterly Journal of Economics* 75, pp.451-467.

Neumann, J. von [1945-6], "A Model of General Economic Equilibrium", *Review of Economic Studies* 13, pp.1-9.

Pigou, A.C. [1920], *The Economics of Welfare*, London: Macmillan.

Pollak, R.A. [1983], "The Theory of the Cost-of-Living Index", *Price Level Measurement: Proceedings From a Conference Sponsored by Statistics Canada.*

__________ [1981], "The Social Cost-of-Living Index", *Journal of Public Economics* 15, pp.311-336.

Samuelson, P.A. [1947], *Foundations of Economic Analysis*, Cambridge, Mass.: Harvard University Press.

__________ [1950], "The Evaluation of Real National Income", *Oxford Economic Papers* 2, pp.1-29.

__________ [1953-4], "Prices of Factors and Goods in General Equilibrium", *Review of Economic Studies* 21, pp.1-20.

——— [1961], "The Evaluation of 'Social Income': Capital Formation and Wealth", in *The Theory of Capital*, F.A. Lutz and D.C. Hague (eds.), London: Macmillan.

——— and S. Swamy [1974], "Invariant Economic Index Numbers and Canonical Duality: Survey and Synthesis", *American Economic Review* 64, pp.566-593.

Sato, K. [1976], "The Meaning and Measurement of the Real Value Added Index", *Review of Economics and Statistics* 58, pp.434-442.

Woodland, A.D. [1978], "On Testing for Weak Separability", *Journal of Econometrics* 8, pp.383-398.

——— [1980], "Direct and Indirect Trade Utility Functions", *Review of Economic Studies* 47, pp.907-926.

SECTION V

Price Behaviour and Business Cycles

Comportement des prix et cycles économiques

THE CONSUMER PRICE INDEX AND SIGNALS OF RECESSION AND RECOVERY

Geoffrey H. Moore
Center for International Business Cycle Research
Columbia University

SUMMARY

A method that uses growth rates in leading and coincident indexes of economic activity in the United States to signal the beginning and end of recessions is described. The results demonstrate the capacity of the system to identify in timely fashion the broad swings in output, employment and the rate of inflation. The U.S. signals also are closely related to business cycles in Canada, as identified by the chronology of business cycle peaks and troughs established by Statistics Canada. It is suggested that development of the signal system based upon Canadian leading and coincident indexes would yield similar results with respect to Canadian business cycles. Comparisons of the U.S. signals with Canadian inflation rates, however, revealed no systematic relationship. This result suggests that in Canada inflation is less closely related to the business cycle than it is in the United States, despite the broad similarity in the movements of the inflation rates in the two countries.

RÉSUMÉ

On décrit une méthode qui permet de déceler le début et la fin des récessions en utilisant les taux de croissance dans les indicateurs avancés et coïncidents de l'activité économique aux États-Unis. Les résultats démontrent que le système permet de reconnaître à temps les grandes variations de la production, de l'emploi et du taux d'inflation. Les signaux aux États-Unis sont par ailleurs étroitement liés aux cycles économiques du Canada, comme le prouve la chronologie des sommets et des creux des cycles économiques établis par Statistique Canada. On laisse entendre que l'élaboration d'un système de signaux fondé sur les indices avancés et coïncidents du Canada donnerait des résultats analogues pour ce qui est des cycles économiques canadiens. Toutefois, la comparaison entre les signaux

aux États-Unis et les taux d'inflation au Canada ne fait ressortir aucun lien systématique. Ce résultat laisse croire qu'au Canada l'inflation et le cycle économique ne sont pas aussi étroitement liés qu'aux États-Unis, malgré la similitude générale des mouvements des taux d'inflation dans les deux pays.

Students of business cycles have long been interested in developing ways of recognizing the onset of recession and recovery as reliably and promptly as possible. This has, naturally enough, been one of the items on the research agenda of the Center for International Business Cycle Research (CIBCR), and we have been making some progress on it. Victor Zarnowitz and I published a paper in January 1982 on Sequential Signals of Recession and Recovery.[1] We showed that the system had some potential not only for identifying. U.S. recessions and recoveries but also for marking off periods of accelerating and decelerating inflation. It is the latter attribute of the signals that I shall concentrate upon here.

The Sequential Signal System

First, however, let me explain briefly how the signal system works. It is based upon growth rates in the leading and coincident indexes published by the U.S. Department of Commerce. The coincident index goes by that name because its highs and lows have roughly coincided with the business cycle turns. It includes four components: industrial production, non-farm employment, real personal income, and total business sales in constant prices. The four components have trended upward over the years, and their average growth rate between 1948 and 1979 was 3.3% per year. This figure gives us one criterion upon which to base the signal system, since the index has generally risen faster than this during business cycle expansions, but slows down when approaching a recession.

The Commerce Department's leading index reaches its turning points in advance of those in the business cycle as well as in the coincident index. It has 12 components, all of which either represent actions taken in anticipation of changes in production, employment, income or sales, or factors that influence such actions. The volume of contracts and orders placed for plant and equipment is an example of the first type, while the level of stock prices is an example of the second. The index also is constructed in such a way that its

long-run growth rate is 3.3% per year, and it too has risen faster than that in recoveries, more slowly when a recession lies ahead.

Both these indexes have declined during every one of the eight U.S. recessions since 1948, and have seldom undergone sustained declines at other times. The leading index did decline on two occasions when no recession ensued, in 1950 – 51 and 1966, but then the coincident index merely hesitated a bit. One is well-advised, therefore, to look at both indexes rather than just the leading index alone. Both indexes are also relatively free of erratic movements, at least as compared with their individual components. But they are not perfectly smooth, and the first figures to be published are also subject to revision. Hence, one must allow for the possibility of frequent reversals.

Bearing these considerations in mind, we first convert the indexes into growth rates, using a special formula. The rate is calculated by obtaining the ratio of the current month's index to the average level of the index for the 12 preceding months, and expressing the result as an annual percentage rate. The advantage of this method, as compared with the ordinary month-to-month change, is that the 12-month average used as the base is less subject to erratic movements than any single month.

During a recession we would expect these growth rates to be negative, since the indexes themselves decline. During recoveries they should be positive. The points at which they turn positive may therefore signal a recovery. At the same time, because of erratic movements or revisions, the growth rate may turn negative again. To meet this problem we adopt two rules. One is that the growth rate must not just go above zero but must go above + 1.0% to be counted as positive. The other is that once this happens any retrogression to a lower rate is not counted as a reversal of the signal unless it goes below – 1.0%.

Since the leading index is expected to turn first, the initial signal is based upon it. The second signal of recovery hinges on a faster advance of the leading index and an upturn in the coincident index. That is, the leading growth rate must exceed 4.3% (1% above the 3.3% trend rate) and the coincident rate must exceed 1.0%. The third signal requires a 4.3% growth rate in both indexes. The same rule about reversals applies throughout.

Signals of the onset of recession are defined similarly. Chart 1 shows how the system has worked in recent years. The growth rates must pass through two bands, one around the 3.3% trend rate, the other around zero. The width of the band, ±1.0%, is the same in both cases. The three signals of peaks are designated P1, P2 and P3 while the trough signals are T1, T2 and T3.

The Record at Business Cycle Peaks and Troughs

How do the signal dates compare with the business cycle peak and trough dates established by the National Bureau of Economic Research? Table 1 lists them all, and shows by how many months the signals led (−) or lagged behind (+). At peaks the P1 signal has always occurred before the peak date, but in three instances (out of 10) the signal proved false: no recession ensued, only a slowdown. P2 occurred sometimes before and sometimes a month or two after the peak, while P3 fell within the range of one to five months after the peak. Two of the three false signals were ruled out by P2, and all three by P3.

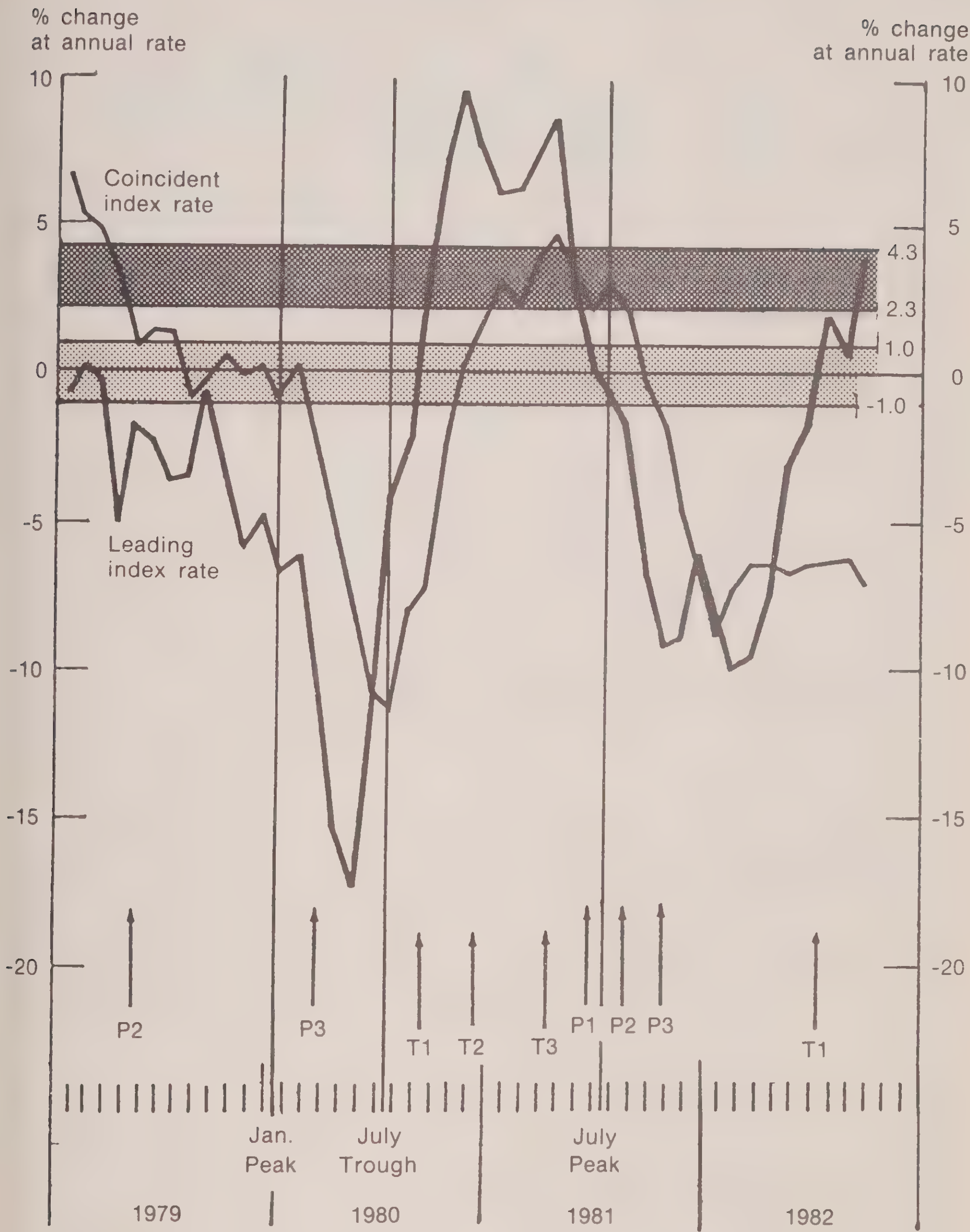
At troughs, T1 has always occurred within three months of the turn, but more often after than before it. T2 has usually occurred five or six months after the turn, and T3 about six or eight months after. There have been no false signals of recovery, although it must be conceded that the record prior to 1977 is based on the revised data now available, not the preliminary data that would be used in practice. Since 1977 we have used the preliminary figures as published each month.

Hence, the system has provided earlier warnings of recessions than of recoveries. For some purposes the tardiness in signalling recovery may not be troublesome (except for politicians!), because at the beginning of a recovery the level of economic activity is low, albeit increasing. Moreover, judging from the record, the use of the first two trough signals alone may yield sufficiently reliable results, reducing the lag by a month or two.

At the present time, the system recorded a first recovery signal (T1) for the current recession in July 1982. At that time the leading index growth rate, having been negative for 12 consecutive months, turned positive, with a rate of 1.9%. A month later, when the August figure was released, the growth rate was 0.7%, but this is not counted as a reversal since

Chart — 1

Recession and Recovery: Sequential Signals



ʹCentre for International Business Cycle Research.

TABLE 1. Signals of Recession and Recovery: Timing at Business Cycle Peaks and Troughs

A. Three Signals of Recession: Timing at Business Cycle Peaks

Business Cycle Peak	P1 First Signal (L<2.3; C>0)	P2 Second Signal (L<-1.0; C<2.3)	P3 Third Signal (L<0; C<-1.0)	Lead (−) or Lag (+), in Months, at Business Cycle Peaks		
				First Signal	Second Signal	Third Signal
11/48	N.A.	N.A.	N.A.	...	...	...
None	3/51	7/51	...	...	...	...
7/54	6/53	8/53	9/53	−1	+1	+2
8/57	1/56	7/56	9/57	−19	−13	+1
4/60	9/59	6/60	9/60	−7	+2	+5
None	5/62	...	...	...	...	...
None	6/66	...	...	...	...	...
12/69	6/69	11/69	4/70	−6	−1	+4
11/73	8/73	1/74	3/74	−3	+2	+4
1/80	11/78	5/79	3/80	−14	−8	+2
7/81	6/81	8/81	10/81	−1	+1	+3
Average	...	...	...	−7	−2	+3

B. Three Signals of Recovery: Timing at Business Cycle Troughs

Business Cycle Trough	T1 First Signal (L>1.0; C<1.0)	T2 Second Signal (L>4.3; C>1.0)	T3 Third Signal (L>4.3; C>4.3)	Lead (−) or Lag (+), in Months, at Business Cycle Troughs		
				First Signal	Second Signal	Third Signal
10/49	8/49	1/50	3/50	−2	+3	+5
5/54	5/54	11/54	12/54	0	+6	+7
4/58	6/58	10/58	11/58	+2	+6	+7
2/61	3/61	6/61	8/61	+1	+4	+6
11/70	11/70	5/71	12/71	0	+6	+13
3/75	6/75	9/75	11/75	+3	+6	+8
7/80	9/80	12/80	4/81	+2	+5	+9
...	7/82	...	...	...	...	...
Average	...	...	...	+1	+5	+8

Note: L and C mean six – month smoothed growth rates in the Commerce Department's leading and coincident indexes, respectively. For 1949-76, based upon revised data as published in 1981. Since 1977, based on preliminary data as published in each successive month.

the rate was still within the lower band. The September figure was 3.9%, which lies within the upper band, and the October, November and December figures (the latest as of this writing) went above the upper band. Meanwhile, the coincident index growth rate continued to be negative, reaching -6.4% in December. Hence, the coincident rate remained well below the lower band, and the T2 signal had therefore not occurred by December 1982. Confirmation of the preliminary signal of recovery was still lacking at the end of the year.

The Signal System and Unemployment

If the signal system is effective in detecting business cycle peaks and troughs, it might be expected to do well in picking out the cyclical swings in unemployment. Chart 2 shows that this expectation is in fact supported. Since the unemployment rate is not a component of the Commerce Department's composite indexes, the signals are statistically independent of that rate. Nevertheless, every interval from P3 to T3, the shaded areas on the chart, was characterized by a rising unemployment rate. Every interval from T3 to P3, except one, was characterized by a declining unemployment rate. The exception was the brief interval between April and October 1981, when unemployment began rising a couple of months before the P3 signal was reached.

Close study of Chart 2 reveals that all of the cyclical highs in unemployment were reached shortly before the T3 signals, while all of the cyclical lows came shortly before the P3 signals. That is to say, the highs in unemployment came in zones marked off by the T1 to T3 signals, while the lows came in the zones marked off by the P1 to P3 signals. Unemployment, in short, obeys the signals.

The Signal System and Inflation

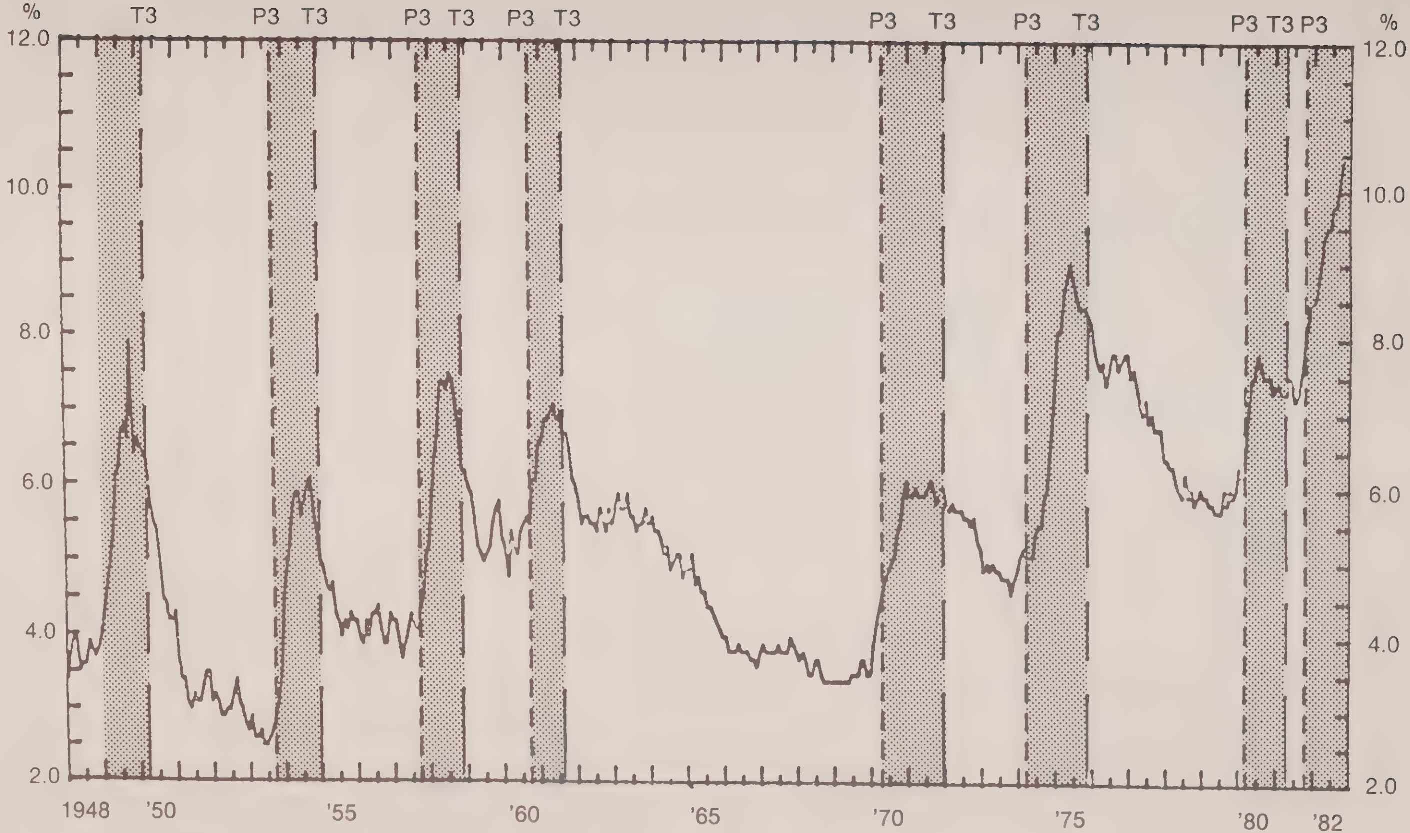
For those who believe in the short-run Phillips Curve, the evidence just presented on unemployment will doubtless be persuasive that the signals have a bearing on the rate of inflation. For those who want the direct evidence we have constructed Chart 3. The inflation rate, based upon the consumer price index, is calculated by the same method that we use to calculate the growth rates in the leading and coincident indexes. That is, the

seasonally adjusted CPI for the current month is expressed as a ratio to the average index during the 12 preceding months, and the result is converted to an annual rate. We call this a six-month smoothed rate, because the interval between the 12-month average and the current month is approximately six months (it is 6.5 months). The rate is smoother than the ordinary six-month change, because the 12-month average is smoother than the single month figure six months ago.

The signals are not as closely related to inflation as they are to unemployment, but they do distinguish periods of rising inflation from periods of disinflation. Each of the intervals from P3 to T3, the shaded areas on the chart, matched a declining phase in the inflation rate. Each of the intervals from T3 to P3 was characterized by a rising inflation rate, with the exception (as with unemployment) of the brief interval in 1981.

Table 2 lays out the record of the inflation rate at each of the signal dates: preliminary, intermediate and final. The average inflation rates at the bottom of the table, covering the period 1949 to 1981, show that at the trough signals the inflation rate was lower than at the peak signals. The lowest average rate came at the second or third trough signal dates, while the highest came at the third peak date. That is to say, one might expect some further decline in inflation between T1 and T2 or T3, and a further rise between P1 and P3. These expectations were realized in nearly all of the individual episodes recorded.

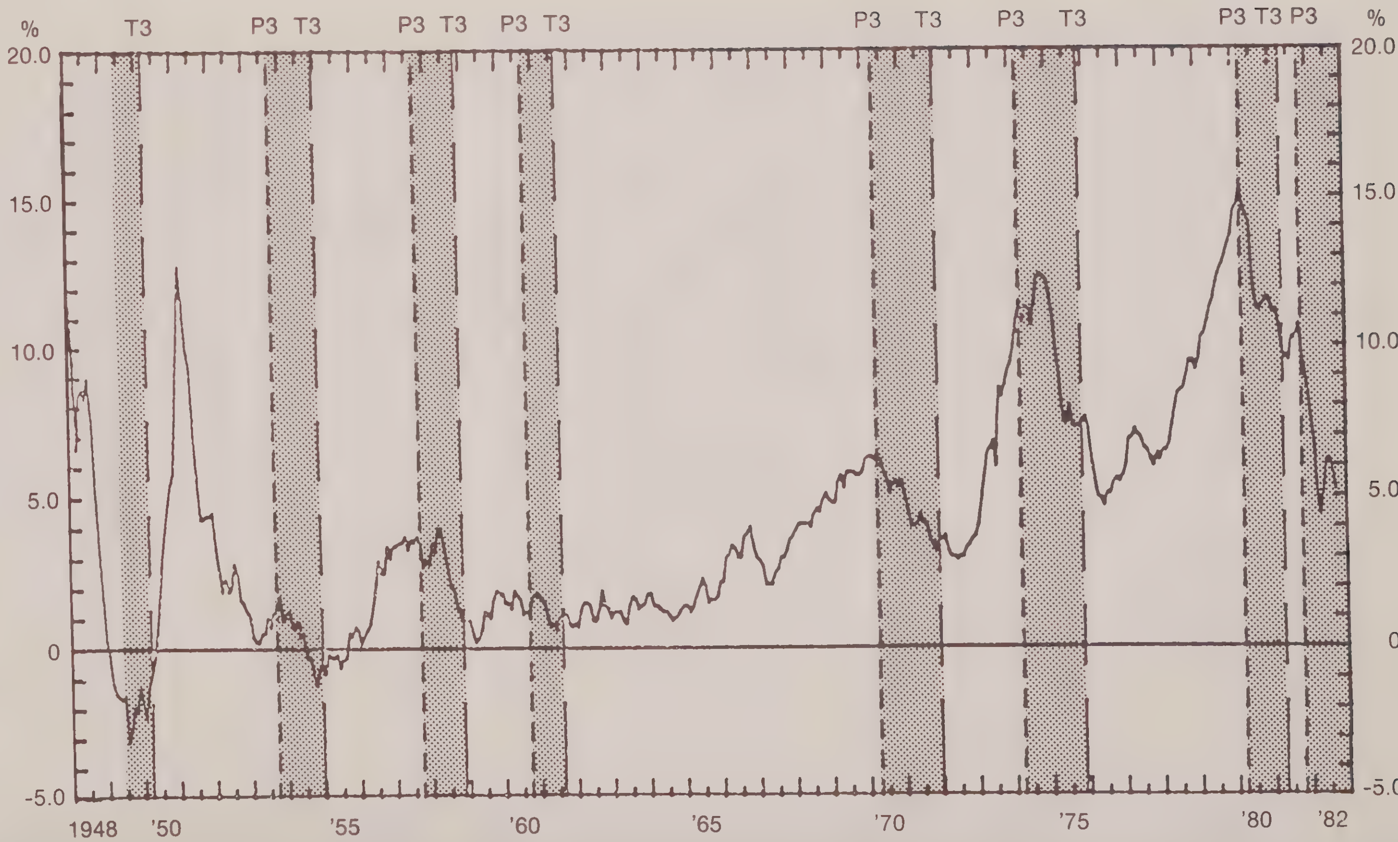
One need not, of course, be a believer in the Phillips Curve to find this pattern of experience credible. The leading index contains many elements that are associated with pressures on prices and costs. A vigorous accumulation of inventories of materials, for example, is likely to show up both in the orders for such goods and in their prices. All three elements, inventories, orders, and materials prices, are represented in the leading index, and increases in materials prices in themselves put upward pressure on many of the prices paid by consumers. To take another example, a boom in housing is likely to boost both orders and prices for furniture, household equipment, lumber and other construction materials, and also to push up mortgage interest rates. Housing permits and orders for consumer goods and materials are in the leading index, while the CPI includes mortgage interest rates as well as prices of houses, furniture and equipment.



P3 = Third signal of business cycle peak (L<0; C<2.3).
T3 = Third signal of business cycle trough (L>4.3; C>4.3).
Center for International Business Cycle Research.

Chart — 3

Sequential Signals of Recession and the U.S. Inflation Rate (CPI)



Shaded areas are business cycle recessions as indicated by the third signal for peaks (P3) and troughs (T3). The CPI rate is the "six-month smoothed percentage change at annual rate", as determined from the ratio of the current month's

TABLE 2. U.S. Inflation Rate (CPI) at Trough and Peak Signal Dates

Period[1]	Trough Signals T1	Trough Signals T2	Trough Signals T3	Peak Signals P1	Peak Signals P2	Peak Signals P3
1949-51	−2.6	−2.3	−0.8	11.4[a]	5.7[a]	...
1953	...	...	...	1.0	1.2	1.3
1954-57	0.5	−0.7	−0.8	0.2	2.9	3.4
1958-60	2.7	1.1	0.9	1.6	1.6	1.1
1961-62	1.0	0.6	0.8	1.4[a]	...	...
1966	...	...	...	3.0[a]	...	...
1969-70	...	...	...	5.8	5.8	6.2
1970-74	5.3	4.1	3.5	8.6	9.9	11.3
1975-80	7.4	7.3	7.5	9.5	10.8	15.3
1980-81	11.2	11.5	10.5	9.5	10.2	9.8
1982	6.2[a]					
Average	3.6	3.1	3.1	5.2	6.1	6.9

[1] For monthly dates of signals during these periods, see Table 1.
[a] Excluded from averages.

Note: The CPI rate is the "six-month smoothed percentage change at the annual rate", as determined from the ratio of the current month's seasonally adjusted index to the average index for the 12 preceding months.

Translating this in terms of the current situation, the inflation rate declined from the last P3 date, October 1981, when it was 9.8%, to the most recent T1 date, July 1982, when it was 6.2%. In percentage points, this 3.6 point decline is about equal to the 3.3 point decline in the averages for these dates, from 6.9% at P3 to 3.6% at T1. Following the T1 signal there has usually been some further decline in inflation, and indeed the decline did continue after July 1982. By December the rate had dropped to 3.4%, and the T2 signal had not yet arrived. Judging from past experience, the next upswing in inflation is not likely to get under way until we have encountered both the T2 and the T3 signals.

These are some of the relationships that help to explain why the sequential signals, based upon the movements of the leading and coincident indexes, are associated with movements of the CPI. To the extent that the relationships persist, the signals may be a valuable tool for monitoring swings in inflation as well as in unemployment. From the point of view of economic policy, it is desirable to have an instrument that has a proven capacity to recognize when unemployment is entering an upswing as well as when inflation is entering a downswing, and vice versa. In practice, many counter-cyclical policies have not had a good record in this respect.

Applying the Signal System to Canada

Since Statistics Canada constructs a composite leading index, one of the prerequisites for the type of signal system described above is readily available. The CIBCR compiles both a leading and coincident index for Canada (as well as for other countries) but so far we have not had the resources to develop a signal system from these indexes. In view of the close relation that has long existed between business cycles in the United States and Canada, however, it might be interesting to apply the U.S. signals to Canadian statistics. To this end, let us first examine the relation between the U.S. signal dates and the Canadian business cycle (Table 3).

What we find is that the U.S. signals bear almost exactly the same relation to the Canadian business cycle turns, as determined by Statistics Canada, as they do to U.S. business cycle turns. The average leads and lags are, in fact, virtually identical (compare Tables 1 and 3). The principal differences are in the 1960s, when several of the U.S. signals have no counterpart in the Canadian business cycle. Also, in 1951 the U.S. signals did not give a definitive indication of the Canadian recession. The latest recession however, was clearly identified. In view of the record, the latest entry in the table, the July 1982 trough signal (T1) should have been as welcome in Ottawa as it was in Washington.

TABLE 3. Comparative Timing of Canadian Business Cycle Turns and U.S. Recession-Recovery Signals

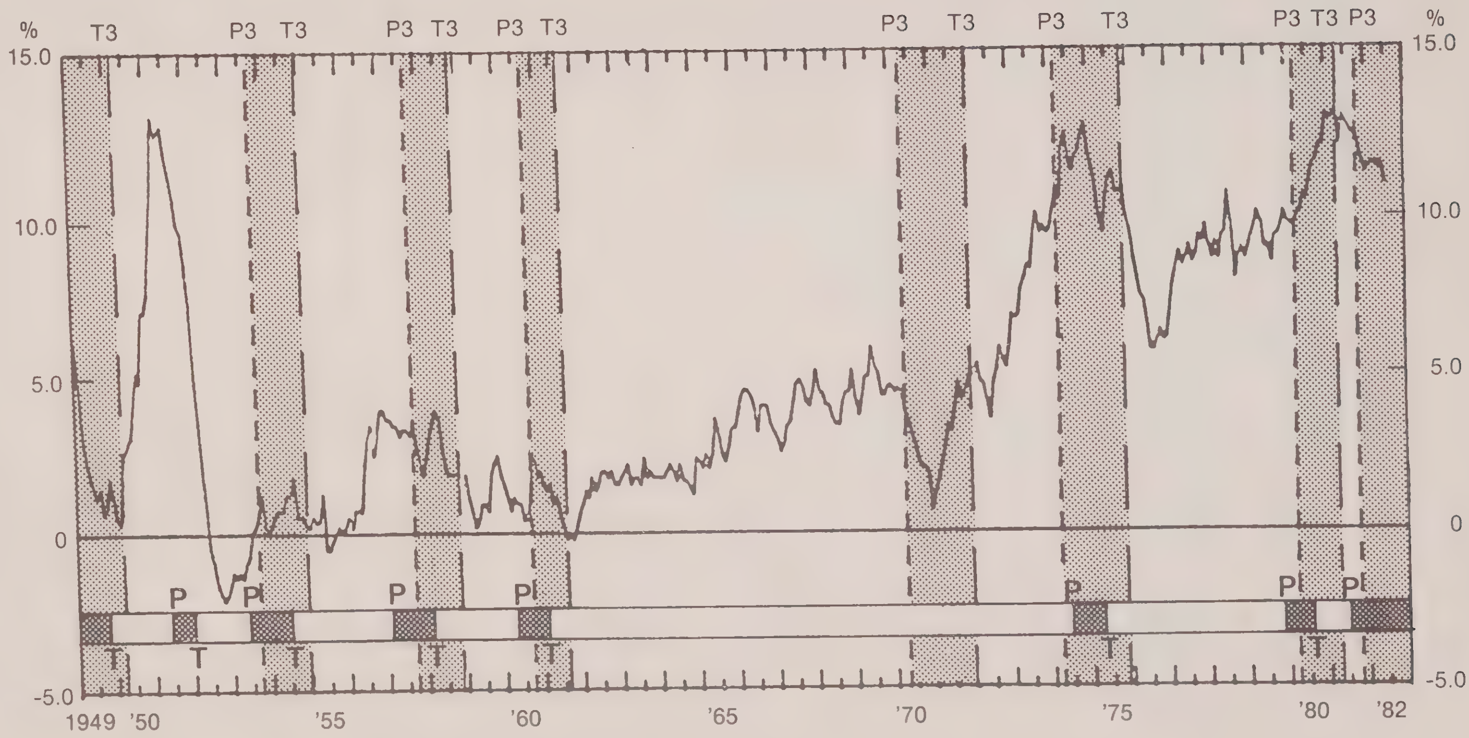
U.S. Signal Dates and Lead(−) or Lag(+), in Months

Canadian Business Cycle		Trough Signals			Peak Signals		
Trough	Peak	T1	T2	T3	P1	P2	P3
10/49		8/49(−2)	1/50(+3)	3/50(+5)			
	5/51				3/51(−2)	7/51(+2)	...
12/51		...	...	...			
	5/53				6/53(+1)	8/53(+3)	9/53(+4)
6/54		5/54(−1)	11/54(+5)	12/54(+6)			
	1/57				1/56(−12)	7/56(−6)	9/57(+8)
1/58		6/58(+5)	10/58(+9)	11/58(+10)			
	3/60				9/59(−6)	6/60(+3)	9/60(+6)
1/61		3/61(+2)	6/61(+5)	8/61(+7)			
	...				5/62	...	...
	...				6/66	...	...
	...				6/69	11/69	4/70
...		11/70	5/71	12/71			
	5/74				8/73(−9)	1/74(−4)	3/74(−2)
3/75		6/75(+3)	9/75(+6)	11/75(+8)			
	10/79				11/78(−11)	5/79(−5)	3/80(+5)
6/80		8/80(+2)	12/80(+6)	4/81(+10)			
	6/81				6/81(0)	8/81(+2)	10/81(+4)
n.a.		7/82					
Average Lead (−) or Lag (+), in months		+2	+6	+8	−6	−1	+4

Note: Canadian business cycle dates, 1951-80, are from Philip Cross, ''The Business Cycle in Canada, 1950-1981, ''*Current Economic Analysis*'', March 1982, Statistics Canada. The 1949 trough date is from W.A. Beckett, ''Indicators of Cyclical Revivals and Recessions in Canada,'' in Geoffrey H. Moore, *Business Cycle Indicators*, National Bureau of Economic Research, 1961, p. 299.

These results suggest that a signal system based upon Canadian leading and coincident indicators would have properties vis-a-vis Canadian business cycles similar to those we have described for the United States. Whether such a system would also differentiate periods of rising and declining inflation in Canada is another question. The U.S. signals seem to show a rather haphazard relationship to the Canadian inflation rate (Chart 4). What this seems to mean is that in Canada inflation is less closely related to the business cycle than it is in the United States. Perhaps we do a better job in exporting our business cycle than we do in exporting our inflation!

Chart — 4
Signals of Recession in the U.S. and the Canadian Inflation Rate (CPI)



The shaded areas are U.S. business cycle recessions as indicated by the third signal for peaks (P3) and troughs (T3). The darker areas are the Canadian business cycle recessions as determined by Statistics Canada. The Canadian CPI rate is the 'six-month smoothed percentage change at annual rate", as determined from the ratio of the current month's seasonally adjusted index to the average index for the twelve preceding months.

Footnotes

[1] Victor Zarnowitz and Geoffrey H. Moore, Sequential Signals of Recession and Recovery, *Journal of Business*, University of Chicago, Vol. 35, No. 1, January 1982. See also Moore, Recession - Recovery: An Early Warning System, *The Morgan Guaranty Survey*, November 1982.

COMMENTS

David Laidler
Department of Economics
University of Western Ontario

Professor Moore's paper is a fine example of the careful descriptive statistical work for which he and his associates, now at Columbia University, and formerly at the National Bureau of Economic Research, are deservedly well-known. Such work is valuable to us in at least two respects. First, the whole system of leading, simultaneous and lagging indicators of the business cycle, developed by the National Bureau, is an invaluable diagnostic tool for anyone concerned with the analysis of contemporary economic policy. Second, and closely related, it is to such work that the economic theorist must look for the so-called "stylised facts" which he uses in his attempts to create analytic models of the business cycle. Without knowledge of a well-defined set of such facts it would be difficult indeed to bring any empirical discipline to bear on the process of theory building and the usefulness of theory as a means of understanding the real world would be diminished considerably.

It is important to realise that descriptive statistical work of the type presented to us here by Professor Moore must be an ongoing enterprise. Though one would hope that a large number of the regularities which past research has uncovered will persist into the future, one can never be certain of such persistence. In particular the Rational Expectations notion suggests that, in principle, once a regular indicator of the course of the cycle has been discovered, agents acting on information provided by that indicator's behaviour might cause its usefulness to be lessened, or at least its significance to be changed. I say "in principle" here, because I gather from Professor Moore that such effects seem to be rare in practice. Nevertheless, so long as the possibility exists of such effects occurring, ongoing research is needed to keep a lookout for them.

To a Canadian audience, by far the most striking and interesting of the results which Professor Moore presents has to do with the behaviour of the Canadian inflation rate relative to that in the United States. It appears, over the period which he has analysed, that although there is a close relationship between fluctuations in real economic variables in Canada and

the United States, the links between the behaviour of price levels are much weaker. Here we have a stylised fact in need of an explanation, and I shall use the remainder of this brief comment to speculate upon the form which such an explanation might take.

There can be no doubt that the fact in question reflects the well-known truth that the short-run Phillips trade-off between inflation and real economic activity in Canada, though it surely exists, is less well determined than it is in the United States. To put the matter this way suggests that one might look for internal differences between the two economies to explain matters, and it is tempting to point to the greater degree of unionisation of the Canadian labour market, and to what is often regarded as that market's greater rigidity, as a starting point for such an explanation. I am rather sceptical about such an explanation, because, even more than Canada, the United Kingdom is alleged to have an inflexible labour market dominated by the activities of trade unions. Nevertheless, as far as the sensitivity of the inflation rate to unemployment is concerned, empirical work seems to show that the U.K. displays *more* flexibility than does the U.S. It is not difficult to rationalise this fact. After all, a powerful union movement may have the capacity to slow down wage adjustments if it so wishes, but it also has the power to speed them up across a wide segment of the economy if it judges that to be in the interests of its members. The empirical evidence suggests that it might be the latter tendency which has dominated British experience, at least over the last 30 years or so. Any explanation of the slow adjustment of inflation to unemployment in Canada along lines under discussion here must explain why the same tendencies have not dominated here.

As an alternative to a labour market rigidity explanation of the facts Professor Moore reports, I prefer one which looks to the nature of the linkages between the United States and Canada. Professor Moore suggests that the United States has been more successful at exporting its business cycle than its inflation rate to this country. Under a flexible exchange rate regime, that is surely just what one would expect. It is a commonplace of international monetary economics that a flexible exchange rate regime cannot insulate a country against real shocks originating abroad. Only to the extent that money illusion prevails can exchange rate variations change the real terms of trade, or offset terms of trade variations originating elsewhere. However, that very same literature tells us that, in the absence of money illusion, and in the presence of a genuinely independent domestic policy towards

the money supply, a flexible rate can provide complete insulation from changes in the behaviour of the general price level in the rest of the world.

Because Professor Moore's period of analysis encompasses both flexible and fixed rate experience, it might have been hoped that his results would throw some light upon this possible explanation of the general nature of the associations among the various series he analyses. Unfortunately, this is not the case. Professor Moore's basic unit of observation is the business cycle phase, and Canada's post-Korean War experience with a formally fixed exchange rate all lies within one long cyclical upswing over the years 1961 – 1970. One cannot learn much from a single observation and if we do conclude that the exchange rate regime accounts for the results we are discussing here, we are doing so on purely *a priori* grounds. That is a dangerous step to take.

As Professor Moore has pointed out to me in conversation, it is generally true for various regional subdivisions of the United States that the real components of their economies correlate more closely with those for the rest of the country than do their price levels. Thus it might be that what we are observing here is nothing more than a reflection of the fact that Canada is, when all is said and done, a region within a highly integrated North American economy. If that did turn out to be the relevant explanation, basic international monetary theory would then force us to conclude that, even though the Canadian - U.S. exchange rate was formally flexible for much of recent history, the monetary authorities were managing its value within rather narrow bounds. Such an hypothesis is far from implausible. Analysis of the type which Professor Moore has presented to us, carried out not just for Canada vis-à-vis the United States, but also for various regions of the United States relative to North America would enable us to see whether Canada's exchange rate flexibility did in fact bestow any extra degree of independence on Canadian policy over the period he has studied. I hope that he and his associates can be encouraged to undertake such an investigation of this interesting question generated by the paper he has presented to us today.

PRICE BEHAVIOR AND ECONOMIC PROSPECTS

Jay W. Forrester
Sloan School of Management
Massachusetts Institute of Technology

SUMMARY

In reaction to unfavorable economic symptoms, governments change regulations, increase expenditures, and alter tax laws. In doing so, governments are running political and economic experiments on full-scale economies. Unfortunately, those responsible for managing countries have not moved aggressively to acquire the means to test policy changes in the laboratory before trying untested policy experiments on entire countries. Better ways are now becoming available for understanding how the structure and policies in an economic system produce its successes and failures. New, fully dynamic computer simulation models can now be constructed for analyzing economic policy alternatives quickly and at low cost before proposed policies are committed to full-scale use.

The System Dynamics National Model is a laboratory representation of a real economy. The Model permits controlled experiments in which a policy in the Model can be changed to determine how behavior would be affected. Model simulations show why policies have so often failed in real economic systems.

Few people suspect the degree to which the puzzling complexities of business cycles, unemployment, depressions and inflation arise from interactions between well-known and well-understood parts of an economic system. The System Dynamics National Model is constructed from policies, organizational structure and physical processes that would be familiar to most businessmen and the model produces the same troubling modes of behavior that have escaped explanation in real life. Actual economic behavior is puzzling, not because of insufficient information about the parts of an economic system, but because, until recently, it has not been possible to show how well-understood parts interact to produce the baffling behavior of the whole system.

The National Model offers a new perspective for understanding inflation. Many anti-inflation policies in recent years have been ineffective because they treated immediate symptoms, rather than underlying causes of inflation. We believe that much of the confusion about inflation arises from failing to distinguish six different processes that cause prices to vary – business-cycles, the economic long wave, economic stresses, inflationary expectations, changing velocity of money and changing money supply.

People in management and politics focus on business-cycle behavior and the accompanying recessions at intervals of three to seven years. But a business cycle does not provide an adequate time horizon for understanding today's economic turbulence. Larger and slower changes are of far greater importance to corporate strategy and to national policy. Economic activity has risen to peaks and then declined into great depressions of the 1830s, 1890s and 1930s. The rise and fall of economic activity between great depressions has come to be called the economic long wave. The long wave, or Kondratieff cycle, which runs its course over some 45 to 60 years, is important in understanding today's economic difficulties. The long wave arises from policies followed in the private sector and causes a rise and fall in capital investment, employment and prices.

Literature on the Kondratieff cycle is filled with debate and conflicting assertions. Economic evidence has been interpreted differently by different observers. Because no theory of the long wave had existed to show how the many aspects of reality could fit into a unified pattern, controversy was unavoidable. But now, the National Model demonstrates a mechanism for generating the long wave. For the first time there is a cohesive theory to explain how an economic pattern spanning a half century can be systematically and internally created within an economy. The long-wave process involves an overbuilding of capital-producing sectors during which they grow beyond the rate of capital production needed for long-term equilibrium. Capital plant is expanded until it exceeds the level justified by the marginal productivity of capital.

Long-wave behavior, as revealed in the National Model, explains many things now happening around the world. Current economic conditions are much like those that the National Model exhibits at a peak of the long wave. At such a long-wave peak one should expect business cycles to become more severe, excess capacity to appear in many industries,

productivity to no longer rise as it had before, return on investment to be declining, prices to rise, debt to increase and unemployment to worsen. Also, at a peak in the long wave, interest rates would be high, the pace of capital investment would slacken and technical innovation would diminish. Such conditions fit the recent past. Similar conditions last occurred in the 1920s at the previous long-wave peak.

We are now past the peak in the long wave and are entering the phase that has usually become a major depression. A depression occurs when long-term growth trends of the past can no longer be sustained. When the trends falter, a powerful reversal follows. For the next 10 or 15 years we should expect unfavorable economic trends. Present conditions indicate that business-cycle recoveries will become progressively weaker and downturns stronger. Along that road, industrial capacity utilization will decline further. Each business-cycle downturn will lead to higher unemployment peaks. Bankruptcies will rise. Developing countries will default on debts and will expropriate the physical assets of multinational corporations. In time, when imbalances have been corrected, when prices and wages have returned to an appropriate relationship to one another, when debts have been liquidated, and when inefficient corporations and social institutions have disappeared, then a new rebuilding can begin.

In describing economic conditions for the next decade, I did not mention inflation. Herein lies the greatest uncertainty. I believe that a physical depression, that is, unemployment and under-utilized capacity, can be accompanied by either deflation or inflation. Whether a country travels the traditional road of deflation or the less well-understood path through runaway inflation depends on the action of government in controlling the money supply. Although the bad news is that industrial economies appear weak, the good news lies in the paradox that our economies have never been stronger. At the end of the building phase, an industrial economy is in the best physical condition it has ever achieved. More capital equipment exists than ever before, productivity is the highest it has ever been, more housing has been constructed, and labor is more readily available. The time following a peak in the long wave should be a golden age, the time toward which society has been striving.

What should industrial countries do as they stand poised at the peak following a capital-investment expansion? What steps should be taken to avoid the historical precedent of drifting into a major depression? Social innovation is needed in several areas – in moving

people into productive and rewarding jobs, in relieving stresses from heavy debt loads, and in handling the clash of forces between inflation and deflation.

RÉSUMÉ

En réaction à des symptômes économiques défavorables, les gouvernements modifient les règlements, augmentent les dépenses et changent les lois fiscales. En agissant ainsi, les gouvernements font des expériences politiques et économiques sur l'ensemble de l'économie. Malheureusement, les personnes responsables de la gestion des pays n'ont pas sérieusement tenté d'obtenir les moyens de tester les changements de politique en laboratoire avant d'entreprendre sur des pays entiers des expériences avec des politiques non testées. On dispose maintenant de meilleurs moyens de comprendre comment les structures et les politiques d'un système économique produisent ses succès et ses échecs. Il est actuellement possible de construire des modèles de simulation informatiques entièrement dynamiques pour analyser les choix de politiques économiques rapidement et à peu de frais avant que les politiques proposées ne soient mises en application sur une économie globale.

Le Modèle national de dynamique systémique est une représentation de laboratoire d'une économie réelle. Le modèle permet des expériences contrôlées dans lesquelles on peut modifier une politique dans le modèle pour déterminer comment le comportement serait affecté. Les simulations du modèle indiquent pourquoi les politiques ont si souvent connu des échecs dans des systèmes économiques réels.

Peu de gens soupçonnent dans quelle mesure les complexités déroutantes des cycles économiques, du chômage, des dépressions et de l'inflation proviennent des interactions entre des éléments biens connus et bien compris d'un système économique. Le Modèle national de dynamique systémique est construit à partir de politiques, d'une structure organisationnelle et de processus physiques qui seraient familiers à la plupart des hommes d'affaires, et le modèle produit les mêmes modes de comportement étranges qui sont demeurés inexpliqués dans la réalité. Le comportement économique réel est étrange, non parce qu'on manque d'information sur les parties d'un système économique, mais parce qu'il n'a pas été possible jusqu'à tout récemment de montrer comment ces parties bien comprises interagissent pour produire le comportement incompréhensible du système global.

Le Modèle national offre une nouvelle perspective pour la compréhension de l'inflation. Plusieurs politiques anti-inflationnistes de ces dernières années sont demeurées inefficaces parce qu'elles s'attaquaient à des symptômes immédiats, plutôt qu'aux causes sous-jacentes de l'inflation. Nous croyons qu'une grande partie de la confusion sur l'inflation provient du défaut de distinguer six processus différents qui causent les variations des prix – les cycles économiques, le mouvement économique long, les tensions économiques, les attentes inflationnistes, les changements de la vélocité de la monnaie et les changements de l'offre de monnaie.

Les gestionnaires et les politiciens se centrent sur le comportement du cycle économiques et des récessions qui l'accompagnent à des intervalles de trois à sept ans. Mais un cycle économiques ne fournit pas un horizon temporel suffisant pour la compréhension des problèmes économiques actuels. Les changements plus amples et plus lents ont beaucoup plus d'importance pour la stratégie des entreprises et pour la politique nationale. L'activité économique a atteint des sommets et a ensuite plongé dans de fortes dépressions dans les années 1830, 1890 et 1930. La montée et la descente de l'activité économique entre les grandes dépressions a été appelée mouvement ondulatoire de longue durée. Ce mouvement ondulatoire, ou cycle Kondratieff, qui dure de 45 à 60 ans, est important pour la compréhension des problèmes économiques actuels. Il provient des politiques du secteur privé et cause la hausse et la baisse des dépenses d'investissement, de l'emploi et des prix.

La littérature sur le cycle Kondradieff est remplie de débats et d'affirmations contradictoires. Des observateurs différents ont interprété différemment les preuves économiques. Étant donné qu'il n'avait existé jusque-là aucune théorie du mouvement ondulatoire de longue durée pour montrer comment les divers aspects de la réalité pouvaient s'intégrer dans un système unifié, la controverse était inévitable. Mais maintenant, le Modèle national présente un moyen de générer un tel mouvement. Pour la première fois, il existe une théorie cohérente qui explique comment on peut systématiquement et de l'intérieur créer dans une économie un comportement économique qui couvre un demi-siècle. Le mouvement ondulatoire de longue durée comprend une période d'hypertrophie des secteurs producteurs de capital au cours de laquelle ceux-ci se développent au-delà du taux de production de capital nécessaire à un équilibre de long terme. On fait des immobilisations jusqu'au moment où celles-ci dépassent le niveau justifié par la productivité marginale du capital.

Le comportement du mouvement ondulatoire de longue durée, comme le révèle le Modèle national, explique beaucoup de choses qui se produisent actuellement dans le monde. Les conditions économiques actuelles sont très semblables à celles que présente le Modèle national à un sommet du mouvement ondulatoire de longue durée. À un tel sommet, on devrait s'attendre à ce que les cycles économiques deviennent plus marqués, qu'un grand nombre de branches d'activité affichent un excédent de capacité, que la productivité cesse de monter comme elle le faisait jusque-là, que le revenu des investissements diminue, que les prix montent, que les dettes augmentent et que le chômage empire. De plus, à un tel sommet, les taux d'intérêt seraient élevés, le rythme des investissements en capital se ralentirait et l'innovation technique diminuerait. Ces conditions s'appliquent bien au passé récent. Ce sont des conditions semblables qui prévalaient dans les années 1920 au précédent sommet du mouvement ondulatoire de longue durée.

Nous avons maintenant passé le sommet du mouvement ondulatoire de longue durée et nous entrons dans la phase qui est généralement devenue une dépression importante. Une dépression se produit lorsque les tendances de la croissance à long terme du passé ne peuvent plus être soutenues. Lorsque ces tendances fléchissent, un violent revirement s'ensuit. Au cours des 10 ou 15 prochaines années, nous devrions nous attendre à des tendances économiques défavorables. Les conditions actuelles indiquent que les reprises du cycle économiques deviendront graduellement plus faibles et les replis plus prononcés. Pendant ce temps, l'utilisation de la capacité industrielle diminuera davantage. Chaque nouvelle baisse du cycle économique entraînera de nouveaux sommets de chômage. Le nombre de faillites augmentera. Les pays en voie de développement cesseront leurs paiements sur leur dette et exproprieront les biens des sociétés multinationales. En fin de compte, lorsque les déséquilibres auront été corrigés, que les prix et les salaires auront retrouvé une relation appropriée, que les dettes auront été remboursées et que les entreprises et les institutions sociales inefficientes auront disparu, alors une nouvelle reconstruction pourra commencer.

Dans ma description des conditions économiques pour la prochaine décennie, je n'ai pas mentionné l'inflation. Là se situe la plus grande incertitude. Je crois qu'une dépression physique, c'est-à-dire le chômage et la sous-utilisation de la capacité, peuvent s'accompagner soit de la déflation soit de l'inflation. Qu'un pays parcoure la voie traditionnelle de la déflation ou le chemin moins bien compris de l'inflation galopante dépend de

l'action de son gouvernement sur le contrôle de la masse monétaire. Bien que ce soit une mauvaise nouvelle que les économies industrielles semblent faibles, il s'en trouve une bonne dans le paradoxe que nos économies n'ont jamais été plus fortes. À la fin de la phase de construction, une économie industrielle se trouve dans la meilleure condition physique qu'elle ait jamais atteinte. Il existe plus de capital que jamais auparavant, la productivité est à son sommet, on a construit plus de logements et la main-d'oeuvre est facilement disponible. La période qui suit un sommet dans le mouvement ondulatoire de longue durée devrait être un âge d'or, l'idéal vers lequel tendait la société.

Que devraient faire les pays industrialisés alors qu'ils se trouvent au sommet qui suit une période d'expansion de l'investissement? Que faudrait-il faire pour éviter le précédent historique de la descente dans une dépression profonde? Il faut faire preuve d'innovation sociale dans plusieurs domaines – replacer les personnes dans des emplois productifs et rémunérateurs, soulager la tension des lourds fardeaux de dettes et faire face à l'affrontement des forces entre l'inflation et la déflation.

The gathering economic storm clouds are darker than any seen in several decades. Economic stresses throughout the world are similar to those described recently in the Nineteenth Annual Review of the Economic Council of Canada: "Economic performance has deteriorated markedly over the past year. The current recession has been both longer and deeper than most observers expected. ... The unemployment rate has moved into the double-digit range and is higher than at any time since the Depression. ... The debt-to-equity ratio of nonfinancial corporations has risen dramatically." Why have the industrial economies come to such a condition? How can we better understand economic behavior so that more effective policies can be chosen?

The debate regarding inflation and unemployment reveals confusion about causes of the current economic stagnation and disagreements about policy recommendations. Inflation has been attributed to a variety of superficial causes including labor unions, insufficient capital investment, and increase in oil prices. Unemployment tends to be blamed without proof on low productivity, foreign imports and monetary policy. Faced with such conflicting opinions, neither governments nor the public have known what to do. Economic theory has provided little persuasive guidance.

In reaction to unfavorable economic symptoms, governments change regulations, increase expenditures and alter tax laws. In doing so, governments are running political and economic experiments on full-scale economies. Because the complexity of economic systems obscures the relationships between policies and their consequences, such experiments fail far too often. And even in the face of such failures, there are few aggressive research programs seeking a better understanding of economic behavior. There are no programs that are funded in proportion to the economic risks inherent in choosing policies without reliably foreseeing their consequences.

By contrast, if a corporate executive were planning a new kind of chemical plant, he would not build it full-scale and use the final system for experimenting with the separate processes and their interactions. Instead, he would first make limited experiments, test with small pilot plants, and use computer simulation models until the processes were better understood.

Is it not strange that for rather simple systems, like oil refineries and space satellites, we use now the best modern methods for analyzing behavior? Yet when faced with the far greater complexity of economic systems we still cling to intuition and debate as the methods for analysis. Surely it is not that the traditional methods for establishing economic policy are serving so well. Newspaper headlines testify to the contrary. Why then is there such disparity between the powerful methods applied in engineering technology and the ineffective methods still used in dealing with economic affairs?

Unfortunately, those responsible for managing countries have not moved aggressively to acquire the means to test policy changes in the laboratory before trying untested policy experiments on entire countries. It is generally recognized that conventional economic theory and the well-known econometric models are not adequate for understanding present economic conditions. Prior to 20 years ago, there were no available alternatives to traditional economic analysis. But, in the last two decades, the methods that have served so well for understanding behavior in engineering technology have been extended until they can encompass the greater complexity and subjectivity of economic systems.

However, governments in most countries still follow analysis methods that have repeatedly been found wanting, without serious'. attempting to develop substantially different alter-

natives for understanding economic behavior. A remarkable disparity in attitude exists between the aggressive methods governments use in dealing with technical challenges in military equipment, compared to the resigned and defeatist approach to policy challenges in economic affairs. A perceived military weakness would immediately trigger multiple research programs along quite different technical approaches in search of the best solution. But, a perceived economic weakness leads to no such aggressive action. Instead, governments in most countries sit immobilized while they receive conflicting recommendations and little useful insight into the causes of economic difficulty. Economic analysis need not be weaker than military technological analysis. Economic analysis should not be weaker, because the international threats from social and economic instability are becoming far greater than are the treats of a military invasion. It is time to attack economic challenges with the same aggressive research and development attitudes that have succeeded so well when applied to complex physical systems.

A. Laboratory Tests of Economic Policies

Better ways are now becoming available for understanding how the structure and policies in an economic system produce its successes and failures. New, fully dynamic computer simulation models can now be constructed for analyzing economic policy alternatives quickly and at low cost before proposed policies are committed to full-scale use. To demonstrate such feasibility, several of us at the School of Management at the Massachusetts Institute of Technology have been developing the System Dynamics National Model. The work is supported by some 40 sponsors. The National Model is a computer simulation model that represents the physical and human processes that interact within an economy to produce business cycles, longer-term economic changes and inflation. The Model shows how the underlying microstructure of an economy generates overall macroeconomic behavior.[1] The National Model provides a simulation laboratory that permits controlled experiments in which corporate and government policies can be changed to determine how economic behavior would be affected. [2,3]

The System Dynamics National Model differs substantially from the more familiar econometric models. The National Model is built up from the operating structure within corporations, rather than from macroeconomic theory.[4] It is derived from management policies as observed in the detailed, practical, working world, rather than from statistical

time series representing aggregate economic behavior. The National Model is intended for evaluating alternative corporate and national policies, rather than for short-term forecasting.

The System Dynamics National Model is a laboratory representation of a real economy. The Model permits controlled experiments in which a policy in the Model can be changed to determine how behavior would be affected. Results are often unexpected and show why such policies have failed in real life. The National Model is sufficiently rich in its structure that it internally generates realistic economic behavior. It does not use the external, exogenous, time-series inputs that must be assumed for important economic variables in the usual econometric models.

The National Model contains internal structures covering a wide range of time behaviors. In the short term are inventory-management and price-setting policies; in the medium term are capital-investment policies, and in the long term will be demographic and environmental changes. By encompassing a diversity of short-term through long-term forces, the National Model can deal with long-range issues of economic growth, resources, energy, and capital investment, as well as with shorter-term dynamics of business cycles and economic stabilization policies. The ability to combine long-term and short-term behavior is necessary for comprehensive policy analysis, because different modes of economic behavior can affect one another, and symptoms of different simultaneous modes can be confused with one another. Furthermore, policies that appear favorable from a short-term analysis are often seen in a more comprehensive framework to lead to cumulative, long-term difficulty.

The Model contains production sectors, labor mobility between sectors, banking with saving and lending, a monetary authority with controls over money and credit, government fiscal operations, and household consumption.

Each production sector reaches down in detail to several factors of production, ordering and inventories for each factor of production, marginal productivities for each factor, balance sheet and profit-and-loss statements, output inventories, delivery delay to determine product availability, production planning, price setting, growth expectations and borrowing.

By spanning from national monetary and fiscal policy down to ordering and accounting details within individual production sectors, the National Model bridges between microeconomic structure and macroeconomic behavior. Just as major behavior modes in the economy develop from deep within its structure, the Model generates the same modes from interactions between elements of the microstructure represented within production, consumption, finance and government.

Few people suspect the degree to which the puzzling complexities of business cycles, unemployment, depressions and inflation arise from interactions between well-known and well-understood parts of an economic system. The System Dynamics National Model is constructed from policies, organizational structure and physical processes that would be familiar to most businessmen, and the Model produces the same troubling modes of behavior that have escaped explanation in real life. Actual economic behavior is puzzling, not because of insufficient information about the parts of an economic system, but because, until recently, it has not been possible to show how well-understood parts interact to produce the baffling behavior of the whole system. The task is less to gather still larger bodies of information about components of an economy, and more to synthesize available information to show how the parts interact to produce the troubling economic symptoms.

The National Model generates many patterns of change that have been observed in real life. The Model exhibits business cycles of three to seven years' duration. It shows Kuznets or construction cycles of 15 to 25 years in length. The National Model produces an economic long wave, or Kondratieff cycle of 45 to 60 years between peaks. It manifests stagflation and reveals the causes of simultaneous unemployment and inflation.

Two aspects of our work should be especially relevant at this time – the multiple causes of change in prices, and the underlying forces now producing economic weakness throughout the industrial world. I will take these up in turn and then interpret their meaning for economic prospects in the next decade.

B. Six Sources of Price Change

First, we will consider six economic processes that cause prices to change. The System Dynamics National Model offers a new perspective for understanding inflation. Many anti-inflation policies in recent years have been ineffective because they treated immediate symptoms, rather than underlying causes of inflation. We believe that much of the confusion about inflation arises from failing to distinguish six different processes that cause prices to vary:

1. business cycles
2. the economic long wave
3. economic stresses
4. inflationary expectations
5. changing velocity of money
6. changing money supply.

These six processes that change prices can exist independently, but their consequences are superimposed in the economic data. The effects of all six appear in prices, wages, production, unemployment, liquidity and interest rates. With the consequences of several sources of price movement intermingled in the data, it has been impossible for statistical analysis to separate important from unimportant processes, or to distinguish causes from symptoms. Only in a laboratory setting, where policies and structure can be controlled, is it possible to disentangle the several processes of price change.

B.1 Change in Prices from Business Cycles

Business cycles are an expansion and contraction of economic activity occurring over intervals of three to seven years. Business cycles contribute rising and falling components to total price change. Due to the business cycle alone, prices would decline during recessions by as much as they rise during upturns. No long-term net change would result from the business-cycle mode by itself. The common perception that, since World War II, prices have "ratcheted" upward during successive business cycles arises because business cycles have been superimposed on long-term inflationary forces, chiefly from monetary causes.

The compensating rise and fall of prices from business cycles is of little consequence to inflation. The business-cycle movements in price are small, and they come and go over relatively short time spans. But even so, business-cycle price movement dominates much of the political and professional debate about inflation. Methods of statistical analysis tend to exaggerate the importance of short-term behavior. Ideas arising from changes in prices, production and employment within business cycles have diverted discussion of inflation away from far more important longer-term processes.

B.2 Change in Prices from the Economic Long Wave

Prices also rise and fall over the 45- to 60-year economic long wave, which will be discussed in more detail later. The long wave, or Kondratieff cycle, which runs its course over some 50 years, is important in understanding today's economic difficulties. The long wave causes a rise and fall in capital investment, debt, employment, innovation, production and prices. [5,6,7]

The amplitude of price movement from the long wave is far larger than from business cycles. Heavy debt loads are created during expansion of the long wave, and new money is created when debt is extended by commercial banks. At the long-wave peak, the added money encourages speculation that pushes up the prices of land and commodities. Fluctuating capital investment and debt help drive the long wave.

Like the business cycle, price movement from the long wave is reversible. Deflation during depressions cancels inflation during booms. But, the long wave is more important than the business cycle because the resulting change in prices is larger and contributes much more to economic disruption. Price change from the long wave can easily be confused with true continuing inflation caused by excessive creation of money by the government.

B.3 Increase in Prices from Economic Stresses

Economic stresses can cause a one-time increase in prices, but not a sustained rise in prices, unless the money supply is increased. The stress processes include most of the alleged causes of inflation except for the other five sources of price change discussed here. In a stress process, something pushes up one or more prices in an economy. Such a stress can

be created by a shortage of some commodity, reduced productivity, declining innovation, union bargaining pressure for higher wages, or resistance to downward movement of prices and wages.

A rise in price created by stress alone is self-limiting. The stress induces counterbalancing forces that reduce other prices. The counterbalancing forces appear as higher unemployment, lower liquidity, less production, larger unsold inventories and higher interest rates. In time, the counterbalancing forces cause price readjustments that relieve the stress.

However, governments no longer wait for the stress modes to correct themselves. Tight liquidity, higher unemployment and rising interest rates are the signals that have in recent decades been taken by governments as indicating the need for more money. In response to such economic stresses, monetary authorities have made more money available.

But more money will not relieve a source of stress. More money temporarily increases liquidity, and allows the original stress-inducing force to push prices still higher. Prices will increase fast enough to absorb any amount of increased money. To increase money in response to symptoms of stress only adds inflation to the other symptoms.

B.4 Increase in Prices from Inflationary Expectations

Once prices are rising, strong forces become established to sustain the rise. Inflation creates the expectation of further inflation. Suppliers raise prices and unions bargain for higher wages, all in anticipation of continuing inflation.

Once established, the circular process of prices rising because wages increase, and wages rising because prices increase, will continue until counterpressures dispel the inflationary expectations. If government increases the money supply in proportion to the rise in prices and wages, liquidity may be maintained and the circular process of rising prices and wages will continue. On the other hand, if money were not increased, then rising prices and wages would require a progressively higher velocity of money, which means falling liquidity. With a constant money supply, liquidity will decline until it exerts enough pressure to counteract the upward spiral of prices and wages. The counteracting forces appear as falling demand for goods, which will put a downward pressure on prices, and rising unemployment, which

will put a downward pressure on wages. The severity and duration of falling demand and higher unemployment will depend on how strongly entrenched are the inflationary expectations.

B.5 Increase in Prices from Rising Velocity of Money

Velocity of money is intimately related to prices and to the ability of a monetary authority to control the price level. Other things being equal, if the velocity of money doubles, so will the price level. An increase in the velocity of money is just as powerful in raising prices as would be a corresponding increase in the amount of money. But we hear almost no discussion of controlling money velocity, although it would be rather easy to do.

The practical significance of the velocity of money lies in its relationship to the size of a bank balance needed to support a particular flow of business transactions. Does a bank balance equal a day's worth of transactions flowing through it, or a week's worth, or is the bank balance equal to the value of a month of transactions? The sustainable flow of financial transactions can be increased either by accumulating larger bank balances or by using existing balances more vigorously – that is, by increasing the velocity of money. Monetary authorities continuously debate the control of money supply while quietly acquiescing to actions that increase money velocity.

Many legal, technical, banking and managerial changes are reducing the necessary size of bank deposits. Electronic banking, computers for corporate data processing, money market funds, credit cards, automatic loans for overdrafts, and relaxation of banking regulations all contribute to increased velocity of money. Most of these could be effectively counteracted by suitable legislation and regulations.

In the United States, from 1965 to 1980, the velocity of money rose at about 3% per year, approximately equal to the average growth in real gross national product. With the velocity of money increasing at the same rate as real GNP, the increase in velocity alone would have been sufficient to support economic growth and stable prices without any increase in the actual money supply. Therefore, all increases in the money supply itself between 1965 and 1980 were fully translated into inflation.

Strong incentives drive efforts to increase the velocity of money. But in the long run such efforts are self-defeating. In managing money, each economic unit perceives that it can purchase a larger share of goods and services by using its money balances harder – by increasing its velocity of money. But when most succeed in doing so, the result is only to raise prices. The higher velocity at the higher prices buys only the same actual goods and services as before. Nothing is accomplished, while the effort to sustain a high velocity of money is actually counterproductive. Inflation absorbs the advantage of higher velocity, and the economic system is left with the added overhead costs of the electronics and personnel needed to maintain the high velocity.

But worse, as velocity increases and becomes more volatile, a monetary authority loses more and more of its control over the effectiveness of money. Because of the multiple forces now converging to increase velocity, a direct legal control of velocity is long overdue. Such control would take the form of setting minimum permissible bank balances in terms of a specified number of days worth of average transactions. Broadly applied, a requirement that bank balances and reserves on those balances be proportional to transaction rates would neutralize the incentives for drawing checks on saving accounts and money market funds and for other techniques for increasing money velocity.

B.6 Increase in Prices from Growing Money Supply

An excess of new money created by the government is the only cause of indefinitely continuing inflation. New money is created when the monetary authority buys government bonds. The bonds represent either past or current government deficits. By such monetizing of government deficits, both the underlying money supply and bank reserves are increased. The additional bank reserves allow commercial banks to create still more money by expanding their loans. In the United States, loans extended by commercial banks are the source of about three-fourths of the total money supply.

Increased money raises prices and therefore the nominal value of assets. The higher value of assets justifies still more borrowing. Such borrowing increases the component of money that comes from bank debt; and prices rise still more as money increases.

In the last 15 years in the United States, money velocity has increased enough to compensate for real growth in the economy. No increase in money supply was needed for economic expansion. But during that time money supply was increased some 2.5 times with a resulting increase in price level by about the same amount.

Except for public political reaction against inflation, there seems to be no counterforce internal to the financial system that will stop continuous money creation by government. Without such political control, there is no limit, except eventual economic and governmental collapse, to how far or fast prices can increase. A continuous increase in money supply will cause a continuous increase in prices. Without such an increase in money, no other forces are capable of causing continuous inflation. In the long run, the price level is determined by the effective supply of money. Stopping inflation depends on restricting the money supply and money velocity sufficiently to hold prices at a stable level.

Two major forces drive excessive creation of money. First is the reluctance of governments to tax fully to cover expenditures. Creating new money avoids raising visible taxes. But the resulting inflation is a hidden tax. The load of government on the private economy is determined entirely by government expenditures, not by direct taxes. If direct taxes do not cover expenditures, then hidden taxes will always make up the difference. Inflation and borrowing both create a hidden tax load on the economy exactly equal to the direct taxes for which they substitute.

The second force leading to inflation is the belief that a restrictive monetary policy produces unemployment. We find in our work that such a belief is largely unjustified. Inflation comes from the governmental money-creation mode of price change. The statistical correlations that suggest a relation between real activity and changes in money supply come from the business-cycle mode. In fact, the direction of causality in the business cycle is the reverse of that generally assumed. Rising business activity and rising inventories cause more borrowing and more money creation, rather than the money creation causing the increased activity. The statistical correlations drawn from business-cycle observations have led to the fallacious belief that action initiated by the monetary authority could directly and significantly affect employment. But the fear of a relationship between tight money and unemployment has stood in the way of controlling inflation.

Temporizing with inflation will continue until the several processes that affect prices are separately understood. Such separation of the mechanisms that cause price change can probably not be done by statistical analysis of historical data. Only with a comprehensive simulation model in the laboratory can the several different structures be identified, their importance established, and relevant policies evaluated.

C. The Economic Long Wave

Now consider the forces that on a worldwide basis are causing unemployment, high real interest rates, excess capital plant and burdensome debts. These symptoms come from long-term economic changes that have been gathering momentum for two decades.

People in management and politics focus on business-cycle behavior and the accompanying recessions at intervals of three to seven years. But a business cycle does not provide an adequate time horizon for understanding today's economic turbulence. Larger and slower changes are of far greater importance to corporate strategy and to national policy.

Articles in the press are drawing more frequent comparisons between present conditions and those in the Great Depression of the 1930s. They could as well compare the present with the Great Depression of the 1890s, or even the Great Depression of the 1830s. There has been a sequence of great depressions. The rise and fall of economic activity between great depressions has come to be called the economic long wave.

The long wave, or Kondratieff cycle, which runs its course over some 45 to 60 years, is important in understanding today's economic difficulties. The long wave arises from policies followed in the private sector and causes a rise and fall in capital investment, employment and prices.

In Western industrial economies, capital investment has been concentrated in periods of economic excitement lasting about three decades. Such periods of active new construction have been interrupted by major depressions. Three such major capital-investment cycles have occurred since 1800. Vigorous economic activity has been terminated by Great Depressions in the 1830s, 1890s and 1930s.

Literature on the Kondratieff cycle is filled with debate and conflicting assertions. Economic evidence has been interpreted differently by different observers. Because no theory of the long wave had existed to show how the many aspects of reality could fit into a unified pattern, controversy was unavoidable.

But now, the National Model demonstrates a mechanism for generating the long wave. For the first time there is a cohesive theory to explain how an economic pattern spanning a half century can be systematically and internally created within an economy. We believe the National Model now provides a theory of how the economic long wave is generated.

The long-wave process involves an overbuilding of capital-producing sectors during which they grow beyond the rate of capital production needed for long-term equilibrium. Capital plant is expanded until it exceeds the level justified by the marginal productivity of capital. At least three forces cause the overexpansion of physical capital – the excitement caused by past growth trends, the varying real interest rates caused by inflation and deflation, and the destabilizing effect from the self-reinforcing demand for capital plant that arises because capital-producing sectors must order their own new capital plant from themselves. Finally, the overexpansion of capital plant is ended by a depression during which excess capital plant is physically worn out and is financially depreciated on the account books until the economic stage has been cleared for a new era of rebuilding.

The process of decline and revival produced by the long wave can be traced by thinking back to the 1920s. In the '20s, prices of assets had been rising. Nominal interest rate was high but, when corrected for rising asset prices, real interest rate was low. As a result of low real interest, physical assets were expanded until investment opportunities in the technology of that time had diminished. Cities were afflicted by excess office space. Land prices had been rising sharply. Heavy debts had been incurred, and the financial system was overextended. Foreign debts had risen beyond the capability of many countries to repay.

Under economic stresses that reached a climax in the late 1920s, previous trends faltered and reversed in the 1930s. The earlier pace of capital investment collapsed. Overcapacity appeared in much of the industrial world. Real interest rates rose to 10% and higher in the early 1930s. Nominal interest was very low but deflation elevated real interest. Prices fell under the pressure of oversupply, and wages were driven down by falling profits and

high unemployment. Developing countries defaulted on their debts and expropriated foreign investment. Bank failures increased. Expansion ceased because agriculture and consumer sectors had excess capacity, so demand fell for construction and capital equipment. For 15 years the industrial economies coasted on the physical capital plant that had been accumulated before 1930. Defaults and bankruptcies by countries, corporations, and individuals cleared out the excess debt load.

By the end of World War II, physical capital at all levels had been depleted. Consumers needed housing and durables. Industry needed factories and machines. Society needed school systems and highways. So a great rebuilding began. Capital sectors drew labor from consumer sectors, thereby producing a tight labor supply and still more incentive for consumer sectors to become more capital intensive. Demand for capital was self-reinforcing: in order to expand, capital sectors themselves required new capital plant, thus creating still further demand on capital sectors. A long, powerful regenerative process drove the expansion of the capital sectors.

Capital plant was rebuilt in a rather brief 20 years from 1945 to 1965. To meet the high demand during those two decades, construction capacity in the capital sectors became far greater than would later be required merely to replace continuing physical depreciation.

Once capital plant had been rebuilt, the capital-construction process did not immediately slow to a sustaining rate. Assume for the sake of illustration that an appropriate level of capital plant was reached in 1965, as indicated since that time by declining return on investment in new capital plant, and by progressively deeper business-cycle recessions. Nevertheless, the capital accumulation process did not stop in 1965 because tremendous momentum had been established in the previous years. Capital investment had become stylish. Labor unions wanted to build more highways and public buildings. Banks had loaned money successfully on earlier construction and aggressively sought new, even though riskier, opportunities to make loans. National monetary authorities had provided credit expansion in the 1950s and 1960s without producing serious inflation, and assumed the process could continue. For more than a decade after 1965, physical capital has been forced into many of the industrialized economies beyond the point where capital earns an adequate return. But such growing investment in the face of falling incentive cannot continue indefinitely.

Eventually, an excess of capital plant suppresses new orders for more capital plant. Such excess capacity on a worldwide basis has appeared in steel production, shipping, diesel-engine manufacture, synthetic textile fibers, chemicals, electronics and automobiles.

Long-wave behavior, as revealed in the National Model, seems to explain many things now happening around the world. Current economic conditions are much like those that the National Model exhibits at a peak of the long wave. At such a long-wave peak one should expect business cycles to become more severe, excess capacity to appear in many industries, productivity to no longer rise as it had before, return on investment to be declining, prices to rise, debt to increase, and unemployment to worsen. Also, at a peak in the long wave, interest rates would be high, the pace of capital investment would slacken, and technical innovation would diminish. Such conditions fit the recent past. Similar conditions last occurred in the 1920s at the previous long-wave peak.

The economic long wave produces large variations in prices and unemployment. At the peak, prices can rise two or more times their depression levels. During depressions, unemployment caused by the long wave can reach 20% or more.

Aggressive expansion of credit, especially during the last part of a capital boom, can make the long wave more severe. Since 1965 the rate of growth in new plant has been slowing. Return on investment has been declining. Productivity has stopped rising, not because there is inadequate capital plant, but because more than adequate plant now exists for each worker. Unlike the 1950s, still more capital plant will now do little to increase productivity. Rather than accept the evidence that capital accumulation had reached its proper goal, industrial countries have tried through easy credit to continue the growth of capital plant. Such generous credit can sustain the boom for a very few years, but at the subsequent cost of a deeper economic downturn than would otherwise have occurred.

A significant part of present inflation arises from the money created by long-wave borrowing. As money is used to bid up the price of assets, those higher-priced assets can support even more borrowing and more money creation. As prices rise, inflation depreciates the value of money. People then try to hold lower money stocks and that causes a higher money velocity. Higher velocity makes the money stock even more powerful in pushing up prices. If monetary authorities had forcefully restricted credit and controlled money

velocity beginning about 1965, the recent peak of the long-wave expansion would have been reduced, and the forthcoming readjustment would have been less difficult than is now likely. But almost no one wanted tight money in 1965. Every element of society supported the actions leading to present difficulties.

We are now past the peak in the long wave and are entering the phase that has usually become a major depression. A depression occurs when long-term growth trends of the past can no longer be sustained. When the trends falter, a powerful reversal follows. Consider real estate prices as an example. Real estate has for two decades been seen as an inflation hedge. Real estate prices rose faster than inflation, so people bought real estate as protection against inflation. Such purchases drove prices still higher. But any prices that rise faster than inflation eventually reach levels far out of balance with other prices. Some farmland is now priced so high that interest on the investment is several times greater than the agricultural value of the land. When, as it must, such an excessive rate of increase in price falters, there remains no economic justification for the heights to which prices have already risen. Prices then fall faster than they rose. When prices start declining, buyers disappear, and prices fall still faster.

A depression period has traditionally started with a rapid deflation in the value of physical assets. That has already been happening. Housing prices in some parts of the United States have fallen 30% and more. Office rental rates are declining. Diamonds and gold have fallen by half.

The economic long wave has persisted through at least three cycles. One often encounters the belief that repetition of the long wave is implausible because of the changes that have occurred in modern economies. But whether or not a significant change has occurred must be judged in terms of a theory of how the long wave is generated. We find that the long wave depends on 1) production being based on the use of capital equipment, 2) processes existing for expanding credit, 3) the life of capital plant being long, some 10 to 40 years, 4) the lifetime of people and 5) the existence of speculative attitudes based on extrapolation of recent trends. Though many aspects of economic structure have changed substantially over the past 200 years, these five fundamental conditions underlying the economic long wave have changed very little. And if one believes that changes in the structure of the economy should have extinguished the long wave, what about the much better known

short-term business cycle? The business cycle has continued for at least 200 years through some 40 cycles. Both the business cycle and the economic long wave have persisted because they are rooted in the fundamental manufacturing and investment processes and the managerial psychology of the private sector. They will yield only to counterintuitive policies. The intuitive, emotional tendencies of managers and politicians, such as more aggressive responses to economic pressures and attempting to extend good times until they turn into overexpansion, result in accentuating the fluctuations, not reducing them.

Some believe that the lessons learned from the Great Depression of the 1930s should prevent a recurrence. But the lessons have been unclear. There is still controversy over the cause of the economic disaster of the '30s. Most economists look on the Great Depression as merely an unusually severe business-cycle downturn. They take a short-term view that does not recognize the decades of speculation, excessive debt creation and overbuilding that preceded the downturn. And even the lessons that were learned are in time forgotten and ignored. This is where the length of human lifetime enters. We would not now have the overbuilding and the risky bank loans if the managerial conservatism following the Great Depression still prevailed. But several generations of management succession had led by 1970 to the go-go managers and bankers who believed they were no longer subject to the lessons of the past.

The economic long wave does not run a separate course in different countries. One should not expect one industrial country to be in the rising phase while another is in the depression phase. Similar systems operating side by side tend to become ''entrained''. That is, even small influences tend to pull them into synchronism. Actually there are very large influences synchronizing economic behavior of countries in the form of foreign trade, money movements and the communication media. So, we see the same pattern of economic symptoms now developing in almost every country. Although each country blames its political leadership, the causes of the economic long wave lie much deeper than political affiliation or economic ideology.

The economic long wave has an important relation to technological innovation. Some people have suggested that the appearance of great new inventions causes each successive capital-construction boom. Indeed, each expansion wave does occur around a new mix

of technologies. However, it seems that the long wave causes a bunching of innovations, rather than that the innovations cause the long wave. We find that the economic long wave of some 45 to 60 years between peaks can exist without technological change. Even with constant technology, capital plant overexpands leading to collapse of capital-producing industries, followed by a 10- or 15-year period during which capital plant is worn out and depreciated before rebuilding begins again. But, the depression periods present a window of opportunity for technological change. The old technology has been overextended and has faltered; the old plant is wearing out; the companies that produced the old capital infrastructure have gone out of business. Technological innovations that have been waiting from 10 to 100 years for their time to come have the opportunity to coalesce into a new social and economic pattern.

The economic long wave has been confusing when observed in real life. The many accompanying symptoms have been taken as coincidences rather than being seen as mutually consistent manifestations of one massive economic syndrome. Only when the long-wave behavior emerges in a comprehensive computer simulation model can the essential driving relationships be appreciated. From such a dynamic representation one can see how the long-wave fluctuation arises from interactions among ordering of capital plant, time-phasing of capital production, self-ordering of capital plant by the capital sectors themselves, influence of growth trends on plans for business expansion, shifts in labor-force participation, rise and fall of housing construction, fluctuations in real interest rate, changes in household saving and credit creation within the banking system. Such comprehensive insights are urgently needed for coping with current economic conditions. Otherwise, governments will continue to shift from one low-leverage policy to another while the inherent economic forces run their course.

D. Economic Prospects

D.1 Stagflation

Economists and political leaders have been surprised by stagflation. Stagflation is the appearance at the same time of unemployment and inflation. A few years ago it was believed on the basis of the so-called Phillips curve relationship that a country could have either

unemployment or inflation but not both simultaneously. The misunderstanding arose from failing to recognize the many different economic mechanisms that can cause unemployment and can change prices.

Recognizing the several different causes of price change helps to explain stagflation. Stagflation occurs because the unemployment arises from behavior within the real, private-business side of an economy, while the inflation comes from government increase in money.

Rising unemployment is now being caused by both the economic long wave and by the economic stress processes. From the long wave, as we now reach the end of the capital-construction boom, unemployment appears first in the capital-producing sectors. Also excess capital plant replaces workers in the goods-producing industries. By the end of the long-wave expansion, housing stocks have been built up, so that less new construction is required. At the same time, the economic stress processes are also increasing unemployment. Take the recent but temporarily reversed oil shortage as an example of an economic stress. When oil prices rose, they increased the average price level. To the extent that the money supply did not correspondingly increase, the higher rate of money flow reduced liquidity. The lower liquidity restrained purchases, lowered demand and reduced employment. The lower liquidity would continue until some prices and wages in the economy had fallen far enough to compensate for the one price that increased.

But in the last three decades, governments have tried to actively intervene to control economic behavior. In response to rising unemployment, high interest rates and reduced liquidity, governments have used monetary and fiscal policies to increase the money supply. However, increased money is very weak in combating unemployment arising from real physical causes within an economy. Instead, increased money leads primarily to bidding up prices of goods and services.

In stagflation, unemployment results from government social policies and from changes occurring in the business and consumer sectors of an economy. The accompanying inflation is a consequence of futile government efforts to stimulate economic activity by increasing the money supply.

D.2 The Future, If We Do Not Change Direction

At the beginning, I referred to gathering economic storm clouds. The metaphor intentionally implies that the impending economic storm has yet to break. Is there time to take shelter? Considerable latitude exists within which to influence the outcome. But, without a better understanding than now exists of the reasons for current economic conditions, the worst scenarios may well prevail. Let me sketch the future toward which we are now headed. Then we can discuss first steps toward changing direction.

For the next 10 or 15 years we should expect unfavorable economic trends. Present conditions indicate that business-cycle recoveries will become progressively weaker and downturns stronger. Along that road, industrial capacity utilization will decine further. Each business-cycle downturn will lead to higher unemployment peaks. Bankruptcies will rise. Developing countries will default on debts and will expropriate the physical assets of multinational corporations. In time, when imbalances have been corrected, when prices and wages have returned to an appropriate relationship to one another, when debts have been liquidated, and when inefficient corporations and social institutions have disappeared, then a new rebuilding can begin.

In describing economic conditions for the next decade, I did not mention inflation. Herein lies the greatest uncertainty. I believe that a physical depression, that is, unemployment and underutilized capacity, can be accompanied by either deflation or inflation. Whether a country travels the traditional road of deflation or the less well-understood path through runaway inflation depends on the action of government in controlling the money supply. If a government attempts to buy its way out of economic stress by creating new money, there could be progressively accelerating inflation. In Germany in the 1920s it has been reported that stores marked up prices four times a day in response to the changing foreign exchange quotations. As inflation accelerates, real economic activity grinds to a halt as more and more of everyone's time is devoted to coping with inflation itself.

On the other hand, if inflation is controlled, the prices of land, houses and corporate equities will be perceived as overpriced compared to wages and goods. A deflation of asset values must follow. I suggest that many industrial countries face odds of five chances in 10 for a runaway inflation, four chances in 10 for deflation, and one chance in 10 for

steering a course between. If one must choose between a rapidly accelerating inflation or deflation, deflation would be preferable. Deflation would run its course sooner than would inflation. Of the two, deflation would lay a more solid foundation sooner for rebuilding the industrial economies. But the great challenge is to steer along that one chance in ten between inflation and deflation. I am not yet ready to suggest how to follow the narrow path between inflation and deflation. To explore that question is on our immediate agenda.

D.3 Favorable Aspects of a Long-Wave Downturn

A major depression is forbidding to contemplate. But it does yield some important economic benefits. It rebalances the economy. It clears away accumulated past excesses.

A depression produces a technological hiatus. Old technologies are given a chance to die out. The economy becomes more receptive to innovative new ideas. Without major depressions the average pace of technological change would be much slower. Some people would prefer a slower pace, others not.

Depressions create enough stress to terminate inefficient institutions and return their people and assets to more productive purposes. Already, we see corporations reducing excessive overhead. Even governments are coming under enough pressure to reduce employment. People in less necessary parts of the service sector, who are adding substantially to the growing economic overhead, will be returned to direct production of food, goods, and housing. Depressions help to keep a society from choking on its corporate and governmental bureaucracies that would otherwise grow unchecked.

Depressions wipe out untenable debt loads. If a country goes through deflation, debts are defaulted. If hyper-inflation is the route, debts may be repaid in nominal money units, but not in purchasing power. In either case, debt that was unwisely accepted by both lender and borrower is eliminated as an economic impediment.

A depression brings assets values, prices, and wages back into alignment. Overpriced land and housing deflate until they once again come within the price range of the average wage earner.

And, a depression terminates unsound economic practices. Consider, for example, loans from the industrial to the developing countries. Banks have been anxious to make loans because weak countries have been willing to pay high interest rates. Why shouldn't such countries pay high interest when they can borrow still more money to pay the interest? A borrower who is ultimately going to default on the debt is willing to promise any interest rate that the lender asks. At the same time, industry and farmers, through their governments, have encouraged banks to make foreign loans. The loans are seen as permitting developing countries to purchase exports from the developed countries. But when debts are defaulted, this curious charade of lending in order to sell goods will be seen as nothing more than a country's having given away the goods. It may be a helpful philanthropy but it is not sound business. A depression clears away pretexts and puts economic activity back on a sounder footing.

But, can we not correct such ailments without the radical surgery of a great depression? I believe there is a basis for hope.

E. Why Should We Not Be Entitled To A Near Utopia?

Although the bad news is that industrial economies appear weak, the good news lies in the paradox that our economies have never been stronger.

Historically, depressions have been times of excess capacity in nearly every economic sector. Depressions occur after capital capacity has been fully rebuilt. When, as now, a peak in the long wave is reached, industrial countries are capable of delivering a higher standard of living than ever before. More housing has already been constructed and is available. People who are no longer needed to build still more capital plant could turn their attention to previously neglected contributions to the quality of life. At the end of the building phase, an industrial economy is in the best physical condition it has ever achieved. The time following the peak in the long wave should be a golden age, the time toward which society has been striving.

At the present time, ample production capacity exists to fill our needs better than at any time in the past. For 30 years industry has been building capital plant, thus increasing

the output per worker. Productivity is now higher than ever before, even if it is no longer continuing to increase. Labor is available to do what the society needs.

The social and economic question is, what happens after making the investment necessary for sustaining a high standard of living? How do we take advantage of the favorable position we have now achieved? We lose sight of the fact that the process of building capital plant should not itself be our goal; our goal should be a high standard of living made possible by having the capital plant in existence.

What should industrial countries do as they stand poised at the peak following a capital-investment expansion? What steps should be taken to avoid the historical precedent of drifting into a major depression? Social innovation is needed in several areas – in moving people into productive and rewarding jobs, in relieving stresses from heavy debt loads, and in handling the clash of forces between inflation and deflation.

As the capital-investment wave ends, people are left in jobs that can no longer be justified, even while diverse unfulfilled tasks exist that could increase the quality of life in a country. Capital-producing sectors will need progressively less employment. Overhead activities as in government, in the service sectors and in manufacturing industries have grown to as much as 90% of the working population and are too heavy an overhead load on the small fraction of people actually producing food and goods. Too many people live on welfare payments, unemployment compensation and retirement income without contributing to the common good. Social innovation must find a way to let more people make a tangible, rewarding contribution to the national quality of life.

Another political challenge lies in the management of debt. Historically, borrowers have been unable to pay back all the debt after a peak in the long wave. Deflation has in the past led to erasing debt by defaults, thereby paying back only a fraction of the amount due. As we move out of a long-wave peak, can debt be repaid? If not, is there a better way to ease the debt burden than by bankruptcies or inflation? Should there be a national action to postpone repayment or to write off part of the debt? A legal innovation must be found to keep debt from dragging national economies into disaster.

The route to the needed social, legal and political innovation lies through a better understanding of present conditions. People are searching for a future they can trust and understand. Understanding is the key. In the last two centuries man has learned to understand technology. During the same time, there has been little progress in truly understanding the behavior of social and economic systems. But now the picture is changing. Methods are becoming available for understanding the complexities of our social and economic institutions.

In considering the next step toward better economic understanding, the presently existing circumstances should be kept in mind. There is no generally accepted explanation for why present economic conditions prevail. Countries with very different economic philosophies are encountering the same difficulties, so the problems cannot be blamed on a particular political party or on special national circumstances. Governments tend to pick low-leverage policies. After such policies fail, each government is replaced by one with equally unfounded and futile programs. The process continues until economic imbalances have corrected themselves. Economic treatments have repeatedly been prescribed before the ailments have been properly diagnosed. The whole process continues in the fog of inadequate understanding. Achieving persuasive understanding must be the first step.

F. Special Committees of Inquiry into Economic Understanding

I suggest that each major nation should appoint a Special Committee of Inquiry into Economic Understanding. The charge to such a committee would not be to recommend policy, or even itself to achieve economic understanding. Instead, each committee should inquire into why there is still so little understanding of economic conditions and their causes. Why is knowledge of economic behavior so much less satisfactory than knowledge of technology? After evaluating the current state of knowledge, each committee should recommend to its government the several most promising and very different avenues to be pursued toward understanding. This should lead to competitive alternative programs following markedly different approaches within a country. Past technological progress has come from multiple parallel efforts in different companies and countries. In economics also, we can expect that breakthroughs in understanding will come from parallel competitive efforts.

I believe that a frontal attack on the economic unknowns can now break through the present confusion and controversy about economic behavior. Past methods of organizing economic research have not been sufficient. Research by individual university faculty members lies below the critical mass necessary for problems of this complexity. It was only through major integrated projects that success was achieved in atomic power, radar, space flight and electronic computers. The economic problems likewise now lend themselves to major well-funded group efforts. Traditional attitudes toward economic research should now be superseded by methods and philosophies that have proven successful outside the social sciences.

Each Special Committee of Inquiry into Economic Understanding should operate in an atmosphere of overriding national priority. A year might be required for its work. Members should be asked to devote undivided full-time attention to the work of the Committee. The chairman of each committee should be not an economist but instead should be an impartial representative of the country in reviewing the shortcomings in economic understanding. Each committee would inquire into why there is so much conflicting opinion between different schools of economic thought. What do they have in common? Wherein lie the fundamental differences? Are economic theories well-founded in reality, or do they represent logical structures built on unrealistic assumptions? To what extent are economic ideas derived from fragments of the entire economic system? Is economic theory so constrained by static analysis that it fails to address the major dynamic changes now confronting the world?

Members of such Committees of Inquiry should be chosen for their concern for economic well-being of the country and their ability to evaluate existing practice and to judge untested proposals. Among committee members, professional economists should be a minority and should be chosen to reflect widely diverse schools of economic thought. The committee for each country would be charged with making an evaluation from outside of the profession of the state of economic understanding that is so crucial to the national future.

From each Special Committee of Inquiry into Economic Understanding would come recommendations for several major projects directed to improving economic understanding. The several projects should represent fundamentally different approaches. Each project would be organized, staffed and funded in the manner that its proponents consider

most appropriate. The government should allocate funding as needed. There should be no hesitation in appropriating tens of millions of dollars to individual projects if needed. Such sums are trivial compared to military budgets. Economic defense should be seen as closely related to military defense. If economic stresses are not relieved, they increase the probability that social pressures will erupt into war.

May I suggest that we take courage to attack our social and economic difficulties with the full skill and daring that we have at our disposal. Now is not the time to accept defeat. The future depends on our willingness to use our capabilities to understand better where we are going. On the basis of better understanding, policies can then be devised that lead to a more favorable future.

References

1. Forrester, Jay W., "An Alternative Approach to Economic Policy: Macrobehavior from Microstructure", *Economic Issues of the Eighties,* edited by Nake M. Kamrany and Richard H. Day, The Johns Hopkins University Press, Baltimore and London, 1979.

2. ________ Nathaniel J. Mass and Charles J. Ryan, "The System Dynamics National Model: Understanding Socio-Economic Behavior and Policy Alternatives", *Technological Forecasting and Social Change,* Vol. 9, 1976.

3. ________ "Business Structure, Economic Cycles, and National Policy", *Futures,* June 1976. Also appeared in *Business Economics,* Vol. 11, No. 1, January 1976.

4. ________ "Information Sources for Modeling the National Economy", *Journal of the American Statistical Association,* Vol. 75, No. 371, September 1980.

5. ________ "Growth Cycles", *De Economist,* Vol. 125, No. 4, Leiden, The Netherlands, 1977.

6. Graham, Alan K. and Peter M. Senge, "A Long-Wave Hypothesis of Innovation", *Technological Forecasting and Social Change,* Vol. 17, 1980.

7. Senge, Peter M., "The Economic Long Wave: A Survey of the Evidence", Working Paper D-3262-1, April 1982, System Dynamics Group, Sloan School of Management, Massachusetts Institute of Technology, Cambridge, Mass. 02139, U.S.A.

SECTION VI

Conclusion

REMARKS CLOSING THE CONFERENCE ON PRICE LEVEL MEASUREMENT

Pour vous fournir une version dans la langue officielle de votre choix, le texte anglais est suivi du texte français (p.1175) dans cette publication.

Martin B. Wilk
Chief Statistician of Canada

On behalf of Statistics Canada, I want to thank all participants in this Conference for contributing to assuring the high quality of our statistical products. The quality and breadth of the research presented, and the productive discussion generated, in these past two and a half days, has fully met my expectations of the professional and societal value of this public Conference. The measurement of price level change, like many other social and economic phenomena, is a very complex task, and one that can benefit from joint review by persons of different points of view and professional background.

We live in rapidly changing times and, correspondingly, statistical agencies must be prepared to review basic concepts and methods of its programs and to make such changes from time to time as may be appropriate. Statistics Canada's sponsorship of this Conference confirms our openness to new ideas and to criticisms and our preparedness to consider alternative ways to go about meeting our responsibilities to provide statistics in the public interest.

Contributors to this Conference have demonstrated a major commitment to professional and social responsibility. I would like to respond to the efforts that went into the research, and the commitment of time and effort it represents, by making some commitments of my own.

First, I want to state again our commitment at Statistics Canada to maintain the highest possible professional standards in carrying out our responsibility to produce reliable, meaningful measures of social characteristics and economic performance, including measures of inflation. Precisely because of the very serious and complex problems facing our socie-

ty, currently, it is essential to maintain the comprehensiveness and the full professional integrity of statistical information.

Second, Statistics Canada will assist Professor Diewert and Professor mMontmarquette to assure early publication of the Conference proceedings. Getting such a volume published and distributed in a timely fashion will give the papers and discussion exposure to a wider audience of professionals and other parties. Hopefully it will also be a stimulus for further research and commentary on critical measurement issues, not only in the academic community but in governmental organizations, private businesses and in research institutions.

Third, Statistics Canada will establish an in-house group of economists and statisticians to study the various contributions to the Conference and assess their potential implications for ongoing statistical procedures, both conceptual and methodological. Wherever there may be scope for refinement and improvement of the existing statistical process, changes will be introduced to our procedures at the earliest possible time.

Fourth, Statistics Canada will establish an Advisory Committee on Price Level Measurement, consisting of individuals from outside Statistics Canada with appropriate professional qualifications and experience, to contribute additional scientific breadth and depth and objectivity in our continuing review of price level measurement concepts, methods and priorities. This committee will be asked to recommend such further studies and research as it may deem desirable and to offer advice on improving the usefulness of price measures. I would expect that, from time to time, the Advisory Committee will issue reports and that these will be available to the public.

I believe this Conference has been of very significant importance not only to price level measurement at Statistics Canada but perhaps even more comprehensively to social and economic statistical measurement processes.

Once again I want to express my sincere gratitude to the researchers and the discussants who presented results and ideas to this Conference, and to the organizers and managers of this event – most especially Professor Diewert and Professor Montmarquette and the other members of the Program Committee. Thank you all for participating.

DISCOURS DE CLÔTURE DE LA CONFÉRENCE SUR LA MESURE DU NIVEAU DES PRIX

To provide you with a version in the official language of your choice, the French text is preceeded by the English text (p.1173) in this publication.

Martin B. Wilk
Statisticien en chef du Canada

Au nom de Statistique Canada, je tiens à remercier tous les participants à cette conférence d'avoir contribué à maintenir la haute qualité de nos produits statistiques. La qualité et la pertinence des documents de recherche présentés au cours de ces deux jours et demi, et le caractère productif des discussions que ces documents ont suscitées, ont comblé tous les espoirs que j'avais mis dans cette conférence publique, tant sur le plan professionnel que social. La mesure du niveau des prix, comme celle de bien d'autres phénomènes sociaux et économiques, est une tâche extrêmement complexe qu'une révision réalisée conjointement par des personnes d'opinions et d'antécédents professionnels différents ne peut que faciliter.

Notre monde évolue rapidement de sorte que les organismes statistiques doivent être prêts à revoir les concepts et les méthodes fondamentales de leurs programmes afin d'y apporter, de temps à autre, les changements nécessaires. En parrainant cette conférence, Statistique Canada a démontré qu'il a l'esprit ouvert aux nouvelles idées et aux critiques, et qu'il est prêt à considérer d'autres façons de fournir des statistiques dans l'intérêt du public.

Les participants à cette conférence ont démontré qu'ils avaient à coeur leurs responsabilités professionnelles et sociales. Vous n'avez ménagé ni votre temps ni vos efforts pour réaliser vos travaux de recherche, de sorte que je voudrais à mon tour prendre certains engagements.

Je voudrais d'abord rappeler que nous nous sommes engagés à maintenir le plus haut niveau possible de professionnalisme dans nos efforts pour produire des mesures fiables et pertinentes des caractéristiques sociales et de la performance économique, y compris

des mesures de l'inflation. L'exhaustivité et la pleine intégrité professionnelle de l'information statistique doivent être maintenues précisément parce que notre société fait face actuellement à des problèmes très sérieux et complexes.

Deuxièmement, Statistique Canada collaborera avec MM. Diewert et Montmarquette pour que les délibérations de la conférence soient publiées promptement. Ainsi, un plus grand nombre de spécialistes et de parties intéressées pourront prendre connaissance des documents et des délibérations de la conférence. On espère ainsi inciter non seulement le milieu universitaire mais également les organismes publics, les entreprises privées et les organismes de recherche à entreprendre d'autres études et à formuler d'autres commentaires sur d'importantes questions liées à la mesure de la variation des prix.

Troisièmement, un groupe de statisticiens et d'économistes de Statistique Canada aura la responsabilité d'examiner les fruits de la conférence et d'en évaluer les implications possibles sur les processus statistiques, tant sur le plan des concepts que des méthodes. Chaque fois qu'il sera possible de rajuster ou d'améliorer le processus statistique actuel, nous modifierons nos procédures le plus rapidement possible.

Quatrièmement, Statistique Canada créera un Comité consultatif de la mesure du niveau des prix composé de personnes de l'extérieur qui présentent les compétences et l'expérience voulues pour ajouter une plus grande dimension scientifique à notre travail continu de révision des concepts, méthodes et priorités de la mesure du niveau des prix, et d'en accroître l'objectivité. Ce comité devra recommander les études et les travaux de recherche supplémentaires qu'il juge nécessaires, et donner des conseils sur les façons d'accroître l'utilité des mesures du niveau des prix. Je m'attends à ce que le Comité consultatif produise de temps à autre des rapports qui seront accessibles au grand public.

J'estime que cette conférence a été un événement important pour la mesure du niveau des prix à Statistique Canada, et plus encore peut-être pour la mesure statistique des phénomènes sociaux et économiques.

Permettez-moi encore une fois d'exprimer ma sincère gratitude aux chercheurs et aux intervenants qui ont présenté leurs travaux et exposé leurs idées à cette conférence, à tous ceux qui ont travaillé à son organisation, et plus particulièrement à MM. Diewert et Montmarquette et aux autres membres du Comité du programme. Merci à tous de votre participation.

INDEX – INDICE